Fauth, Benjamin; Decristan, Jasmin; Decker, Anna-Theresia; Büttner, Gerhard; Hardy, Ilonca; Klieme, Eckhard; Kunter, Mareike

The effects of teacher competence on student outcomes in elementary science education. The mediating role of teaching quality

formal und inhaltlich überarbeitete Version der Originalveröffentlichung in:
formally and content revised edition of the original source in:
Teaching and teacher education 86 (2019) 102882, 51 S.

Bitte verwenden Sie beim Zitieren folgende URN /
Please use the following URN for citation:
urn:nbn:de:0111-pedocs-191491 - *http://nbn-resolving.org/urn:nbn:de:0111-pedocs-191491*
DOI: 10.1016/j.tate.2019.102882 - *http://dx.doi.org/10.1016/j.tate.2019.102882*

The Effects of Teacher Competence on Student Outcomes in Elementary Science Education: The Mediating Role of Teaching Quality

Benjamin Fauth[a,b], Jasmin Decristan[b,c], Anna Decker[b,d], Gerhard Büttner[b,d], Ilonca Hardy[b,e], Eckhard Klieme[b,f], & Mareike Kunter[b,d]

[a] Hector Research Institute of Education Sciences and Psychology, University of Tübingen, Europastraße 6, 72072 Tübingen, Germany

[b] Center for Individual Development and Adaptive Education of Children at Risk (IDeA), Frankfurt, Germany

[c] University of Wuppertal, Rainer-Gruenter-Straße 21, 42119 Wuppertal, Germany, decristan@uni-wuppertal.de

[d] Department of Psychology, Goethe University Frankfurt, Theodor-W.-Adorno-Platz 6, 60629 Frankfurt, Germany, decker@paed.psych.uni-frankfurt.de, buettner@paed.psych.uni-frankfurt.de, kunter@paed.psych.uni-frankfurt.de

[e] Institute of Educational Sciences, Goethe University Frankfurt, Theodor-W.-Adorno-Platz 6, 60629 Frankfurt, Germany, hardy@em.uni-frankfurt.de

[f] German Institute for International Educational Research Frankfurt, Schloßstr. 29, 60486 Frankfurt, Germany, klieme@dipf.de

Author Note

This research was funded by the Hessian initiative for the development of scientific and economic excellence (LOEWE).

Correspondence concerning this article should be addressed to Benjamin Fauth, Hector Research Institute of Education Sciences and Psychology, University of Tübingen, Europastraße 6, 72072 Tübingen, Germany. Email: benjamin.fauth@uni-tuebingen.de

The Effects of Teacher Competence on Student Outcomes in Elementary Science Education: The Mediating Role of Teaching Quality

Abstract

In this contribution, we investigate links between teacher competence, teaching quality, and student outcomes in elementary science education. Students' conceptual understanding and interest were measured during two teaching units in a pre-post design (1,070 students, 54 classes). Results show that teacher competence (pedagogical content knowledge, self-efficacy, and teaching enthusiasm) was positively related to students' interest; self-efficacy was positively related to student achievement. Three dimensions of teaching quality (cognitive activation, supportive climate, and classroom management), which refer to the actual teacher-student-interactions in the classroom, mediated these relationships. These results help illuminate the mechanisms behind the effects of teachers on student outcomes.

Keywords: teacher competence, teaching quality, science education, elementary school

The Effects of Teacher Competence on Student Outcomes in Elementary Science Education: The Mediating Role of Teaching Quality

Good teachers can make a difference in their students' progress (Rivkin, Hanushek, & Kain, 2005). Meta-analytic results show that teachers are an important source of variance in students' development in schools (Hattie, 2009). As a consequence, recent research has set out to determine which specific aspects of teachers' professional competence matter for student development. *Teacher competence* is conceptualized as a framework that describes the specific personal qualities that teachers need to meet the high demands of their profession. The concept covers cognitive as well as motivational variables (Baumert & Kunter, 2013). For example, good teachers should have a profound knowledge of tasks and instructional strategies that foster students' conceptual understanding (i.e., pedagogical content knowledge; Sadler, Sonnert, Coyle, Cook-Smith, & Miller, 2013). They should also exhibit a certain degree of motivation to really be able to concentrate on the challenges of everyday classroom instruction (e.g., enthusiasm for teaching; Keller, Goetz, Becker, Morger, & Hensley, 2014).

Recent studies have fruitfully distinguished these aspects of teacher competence from characteristics that reflect the actual practice of teaching in the classroom (Rimm-Kaufman & Hamre, 2010; Zee & Koomen, 2016). The latter has also been referred to as *teaching quality* and has convincingly been described as having three basic dimensions: effective classroom management, supportive classroom climate, and the potential for cognitive activation (Klieme, Pauli, & Reusser, 2009; Pianta & Hamre, 2009). Teacher competence refers to the teacher's personal characteristics (e.g., knowledge and motivation). In contrast, teaching quality refers to teachers' actual behavior and teacher-student interactions in the classroom (Rimm-Kaufman & Hamre, 2010). Thus, while teacher competence should be related to teaching quality, the two are not interchangeable.

The basic theoretical assumption of the present study is that teacher competence is positively related to teaching quality, which in turn has an effect on student outcomes (Kunter, Klusmann, Baumert, Richter, Voss, & Hachfeld, 2013). This basic idea forms the foundation of several recent studies on different aspects of professional competence such as teachers' knowledge (Baumert et al., 2010; Förtsch, Werner, von Kotzebue, & Neuhaus, 2016; Keller, Neumann, & Fischer, 2016), professional beliefs (Kleickmann, Vehmeyer, & Möller, 2010; Kunter et al., 2013; Staub & Stern, 2002), enthusiasm for teaching (Frenzel, Goetz, Lüdtke, Pek un, & Sutton, 2009; Keller et al., 2014; Kunter et al., 2008), and self-efficacy (Guo, Connor, Yang, Roehrig, & Morrison, 2012; Zee & Koomen, 2016).

The thorough examination of teaching quality can illuminate the processes that underlie the relations between teacher competence and student development (Rimm-Kaufman & Hamre, 2010). More specifically, the effect of teachers' personal characteristics on student outcomes might be mediated by teachers' classroom behaviors and teacher-student interactions in the classroom (i.e., teaching quality). Teacher competence, on the other hand, may serve as an important lever that can be used to improve the quality of teaching and student outcomes, for instance, in professional development programs (Kleickmann, Tröbs , Jonen, Vehmeyer, & Möller, 2016).

However, especially within the domain of elementary science education, not many studies have systematically examined the links between teacher competence, teaching quality, and student outcomes (Kleickmann, Vehmeyer, & Möller, 2010; Lange, Kleickmann, Tröbs , & Möller, 2012). In science education, the relation between teacher competence and instructional settings is complex: Reform attempts in many countries focus on inquiry-based learning (Furtak, Seidel, Iverson, & Briggs, 2012), which is considered crucial for fostering students' conceptual understanding of science phenomena. Among researchers, there is growing consensus that science learning should be regarded as an activity of sense-making that can be promoted through practices such as posing research questions, conducting experiments, and discussing the implications (Odden & Russ, 2019). The promotion of science literacy places high demands on science teachers in general (Duschl & Bybee, 2014), and the implementation of inquiry-based learning in the classroom is particularly challenging (Krämer, Nessler & Schlüter, 2015). However, in Germany, as in many other countries, elementary school teachers are generalists who often do not have an academic background in a science-related subject (Brobst, Markworth, Tasker, & Ohana, 2017). Elementary science teachers are often hesitant to teach science which is probably due to their limited pedagogical content knowledge and low self-efficacy (Appleton, 2008; Johnston & Ahtee, 2006; Rice, 2005). Similar to elementary school teachers, early childhood teachers also report that they expect to fail in teaching science to young children (Greenfield et al., 2009), a fact that may be related to their lack of specific university training or professional development with regard to science teaching (Hope et al., 2017). Overall, teaching science in a manner conducive to young students' learning seems to be a real challenge for elementary school teachers (Appleton, 2008; Kleickmann et al., 2016). Accordingly, examining the antecedents of effective science teaching can offer valuable insights into how to promote educational quality and thus effective student learning. Within this domain, we sought to identify specific, measurable aspects of professional competence that are exhibited by successful teachers. How are these aspects of teacher competence in elementary school related to student outcomes? Which aspects are related to high-quality classroom instruction? Does teaching quality serve as a mediator of the relation between teacher competence and student outcomes in elementary science education? It is still an open question whether or not the relationships reported in the literature will be comparable in the domain of elementary science education.

Theoretical Framework

Classroom instruction is the "core business" of teachers (Baumert & Kunter, 2013). Research on teaching quality provides general principles of high-quality learning environments that have also been applied to science education (Pianta & Hamre, 2009). In the following, we will give an overview on the literature about teaching quality and the different aspects of competence that teachers need to provide high quality instruction. In these sections, we will particularly focus on the context of elementary science education.

Teaching Quality

In recent theoretical frameworks of teaching quality, three basic dimensions of teaching quality that are crucial for student learning and motivation have emerged in different studies: cognitive activation, supportive climate, and classroom management (Baumert et al., 2010; Authors, 2014b; Klieme, Pauli, & Reusser, 2009). These dimensions are very similar to the three domains conceptualized in the Classroom Assessment Scoring System (instructional support, emotional support, and classroom organization; Pianta & Hamre, 2009).

Cognitive activation includes challenging tasks, the exploration of concepts, ideas, and prior knowledge, and Socratic dialogue practice as key features (Lipowsky et al., 2009). These classroom practices should foster students' cognitive engagement, which should in turn lead to elaborated knowledge (Klieme et al., 2009). Cognitive activation in science education works by challenging students' preconceptions through a teacher-guided classroom discourse practice as well as by presenting challenging tasks, observations, and experiments that help students to cognitively engage with the learning content (Leuchter, Saalbach, & Hardy, 2014). As research on science education stresses the importance of conceptual change in students (Vosniadou, 2013; Wandersee, Mintzes, & Novak, 1994), cognitive activation in instruction may be considered an important condition for students' cognitive restructuring of science concepts.

Supportive climate (also referred to as *teacher support*; Lazarides, Gaspard & Dicke, 2018) covers specific aspects of teacher-student interactions such as individual, positive, and constructive teacher feedback, a positive approach to student errors and misconceptions, and caring teacher behavior (Brophy, 2000; Klieme et al., 2009). Positive student-teacher interactions have been conceptualized in different theoretical approaches such as attachment theory, sociocultural perspectives, and self-determination theory (Davis, 2003). Self-determination theory suggests that three basic intrinsic needs are associated with human motivation: social relatedness, autonomy, and competence (Deci & Ryan, 2000; Kunter, Baumert, & Köller, 2007). Supportive climate is a vital element of elementary science education because it stresses the role of teacher feedback for students' construction of knowledge, sensemaking, and conceptual change (Authors, 2014b).

Classroom management is a well-known concept in educational research (e.g., Kounin, 1970) that focuses on classroom rules and procedures, coping with disruptions, and smooth transitions. Effective classroom management provides time on task, which can be seen as a necessary precondition to be actively engaged in learning (Emmer & Stough, 2001). The positive effects of well-managed classrooms on student achievement are well-documented in education research (Lipowsky et al., 2009; Seidel & Shavelson, 2007; Wang, Haertel, & Walberg, 1993). Classroom management is of particular importance in inquiry-based science education, where students work a lot on their own experiments in small group settings. These nontraditional classroom settings in elementary science bear special challenges for a teacher's classroom management.

In summary, cognitive activation and classroom management, in particular, have been shown to predict cognitive student outcomes (Kyriakides, Christoforou, & Charalambous, 2013; Lipowsky et al., 2009; Seidel & Shavelson, 2007), whereas a supportive climate was found to be especially connected to students' motivational and interest development (Authors, 2014b; Kunter, Klusmann, et al., 2013).

Teachers' Professional Competence

In recent years, research has made progress not only in identifying successful teachers (e.g., in value added models; McCaffrey, Lockwood, Koretz, Louis, & Hamilton, 2004), but also in describing specific aspects of teacher competence that can be directly measured via teacher surveys or standardized tests (Gitomer & Zisk, 2015). Just as we can for student outcomes, we can distinguish between aspects of teacher competence that are more cognitive versus more motivational in nature (Kunter, Kleickmann et al., 2013). Cognitive constructs comprise teachers' professional knowledge (Shulman, 1986) and beliefs (Fives & Buehl, 2012; Fives & Gill, 2015; Mansour, 2009), both of which have been widely discussed, especially in the field of science education. Motivational constructs in the field of teacher competence cover aspects such as self-efficacy (Klassen, Tze, Betts, & Gordon, 2011; Zee & Koomen, 2016) and teaching enthusiasm (Kunter et al., 2011). It seems that motivational variables are as important for competent teachers as the classic knowledge areas are (Richardson, Karabenick, & Watt, 2014). The following sections provide a brief overview of the most prominent cognitive and motivational constructs as well as their relations to teaching quality and student outcomes.

Cognitive aspects: knowledge and beliefs. Profession-specific *knowledge* is traditionally regarded as a key factor for teachers' vocational success (Anderson, Blumenfeld, Pintrich, Clark, Marx, & Peterson, 1995; Hill, Rowan, & Ball, 2005). Shulman (1998) and Bromme (2001) made a widely acknowledged distinction between *content knowledge* (CK), which refers to a comprehensive understanding of the topics taught, and *pedagogical content knowledge* (PCK), which refers to ways of making this content knowledge accessible to students (Depaepe, Verschaffel, & Kelchtermans, 2013; Krauss et al., 2008). In particular, PCK has been found to be related to student achievement (Depaepe, et al., 2013; Förtsch et al., 2016; Keller et al., 2016; Mahler, Großschedl, & Harms, 2017; Sadler, Sonnert, Coyle, Cook-Smith, & Miller, 2013). At the same time, because teacher education programs are rather general, it is particularly challenging for elementary school teachers to build up knowledge on how to teach science (PCK) as well as to actually understand science phenomena (CK) (Appleton, 2008; Brobst et al., 2017, Gomez-Zwiep, 2008).

Compared to CK, PCK is more proximal to classroom instruction. It comprises three domains that are directly important to good teaching (Kleickmann et al., 2013; Park & Oliver, 2008): (1) knowledge on tasks that foster student understanding, (2) knowledge on good explanations and instructional strategies, and (3) knowledge on students' preconceptions, misconceptions, and the typical difficulties they encounter during learning. Thus, PCK is closely related to the concept of "subject-matter knowledge for teaching" (Hill et al., 2005, p. 372), which does not focus on content knowledge per se but on "how this knowledge is used in classrooms" (Hill et al., 2005, p. 376).

Teachers with higher PCK are better able to implement challenging tasks at a high cognitive level (cognitive activation; Förtsch et al., 2016) and they provide more individual learning support according to student ratings (Baumert et al., 2010). In science education, teacher PCK is relevant for knowing students' typical misconceptions and difficulties in content-related knowledge construction; it is also relevant for knowing the specific requirements for the acquisition of scientific knowledge (Magnusson, Krajcek, & Borko, 1999). It also pertains to teachers' knowledge of instructional strategies, representations, use of models, experiments, and observations in science education (Meschede, Fiebranz, Möller, & Steffensky, 2017).

It can be assumed that teachers' knowledge of good tasks and teaching strategies is more related to cognitive activation, and their knowledge of students' typical misconceptions helps them to individually support students by addressing these misconceptions. As in science learning, students often hold prior conceptions that are not consistent with scientific views, and they need to differentiate, integrate, and restructure their knowledge over time (Schneider, Vamvakoussi, & van Dooren, 2012), teachers' PCK is especially important. Yet, there are few studies addressing in-service teachers' PCK in elementary science (Meschede, Fiebranz, Möller, & Steffensky, 2017).

In the few studies that have been conducted, it has been found that teachers with greater PCK tend to provide a higher degree of individual support, which might also be responsible for the effects of PCK on students' motivation in elementary science education (Lange et al., 2012). This can be explained by the fact that individual student support has been found to promote student motivation (Kunter et al., 2013; Lazarides et al., 2018).

In the field of teacher *beliefs*, researchers have made a distinction between a transmission orientation and a constructivist orientation (Mansour, 2009; Tsai, 2002; Voss, Kleickmann, Kunter, & Hachfeld, 2013). The first orientation goes along with beliefs about teaching as a direct transmission of knowledge from the teacher to the student. The student's role is thus conceptualized as a "knowledge recipient." The constructivist orientation goes along with beliefs that students process new learning content by building on preconceptions and that new knowledge has to be actively constructed by interacting with this pre-existing knowledge (Dubberke, Kunter, McElvany, Brunner, & Baumert, 2008; Staub & Stern, 2002; Voss et al., 2013). The student's role is thus conceptualized as an active "knowledge constructor."

Constructivist beliefs play a particularly important role in elementary science education, where dealing with students' preconceptions and misconceptions is regarded as a key factor for insightful learning (Jones & Carter, 2007; Kleickmann et al., 2016; Vosniadou, 2013). Elementary science teachers often tend to equate science teaching with hands-on activities (Mayer, 2004), thus not paying attention to the role played by students' preconceptions in learning. Research indicates that transmission beliefs can have negative effects on student achievement (Kleickmann et al., 2016), while constructivist beliefs can have positive effects (Voss et al., 2013).

The results of research on the significance of teacher beliefs for the quality of their instruction have been mixed (Fives & Buehl, 2012; Voss et al., 2013). Staub and Stern (2002) reported that science teachers with a stronger constructivist orientation frequently applied tasks with a greater potential for cognitive activation. From a theoretical standpoint, the question that arises is whether and how teachers' beliefs become relevant for their activities in the classroom. Teachers who acknowledge the constructivist nature of learning processes might cognitively activate students during classroom discourse (Voss et al., 2013) and take more care in handling the individual levels of students' conceptual understanding, thus providing more individual support (Dubberke et al., 2008). However, in an analysis that considered several aspects of teacher competence simultaneously, teachers' constructivist beliefs were related to only classroom management but not to any other dimension, and even more surprisingly, this relation was negative (Kunter et al., 2013).

Motivational aspects: self-efficacy and enthusiasm. Research on *self-efficacy* has a long history in psychological research and has led to numerous results in a variety of fields. In Bandura's (1995) social cognitive theory, self-efficacy is defined as "beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments" (p. 3; see also Tschannen-Moran, Woolfolk Hoy, & Hoy, 1998). Teachers' self-efficacy can be defined as the self-perception of competence to perform well in their job, including the management of potentially challenging situations in everyday school practice (Bandura, 1995; Guo, Dynia, Yeager Pelatti, & Justice, 2014; Schmitz & Schwarzer, 2000; Zee & Koomen, 2016). Teacher self-efficacy is related to greater effort invested in teaching, higher degrees of intrinsic motivation, and a greater openness to new ideas and teaching methods (Tschannen-Moran & Woolfolk Hoy, 2001). However, the link between teacher self-efficacy and student outcomes is not as well established as one would expect: In Klassen et al.'s (2011) seminal efficacy review, only two out of 218 studies between 1998 and 2009 focused on this link. A recent meta-analysis revealed only a small effect on student achievement (Klassen & Tze, 2014). According to Klassen et al. (2011), further research on this link is needed—particularly with strong research designs.

Correlational relations between teacher self-efficacy and teaching quality in the classroom have also been documented in the literature (Klassen & Tze, 2014; Zee & Koomen, 2016). Considering the way self-efficacy is conceptualized, positive relations with all three of the above-described basic dimensions can be expected. Teachers with high self-efficacy should be better able to cope with the demands of high-quality classroom instruction reflected by the dimensions of cognitive activation, supportive climate, and classroom management (Zee & Koomen, 2016). Given the special situation of an elementary school teacher who is teaching science, we assume that teacher self-efficacy will be particularly important (Appleton, 2008).

Whereas self-efficacy research focuses on the part of motivation involving expectancies (“I am able to accomplish this”), *teaching enthusiasm* refers to the intrinsic value of teaching activities (“Doing this is fun for me”). Enthusiasm was identified long ago as one of the most important characteristics of good teachers (Brophy & Good, 1986). However, in recent years, researchers have been working on establishing a clear definition of teacher enthusiasm. Drawing on interest and self-determination theory, Kunter et al. (2008) defined enthusiasm as the trait-like “affective component of teacher motivation” (p. 470; see also Keller et al., 2014), which is reflected in enjoyment, excitement, and pleasure during professional activities. The authors distinguished between two dimensions of teacher enthusiasm: enthusiasm for the subject (e.g., “I am still enthusiastic about the subject of mathematics”) and enthusiasm for teaching (e.g., “I really enjoy teaching mathematics in this class”).

Kunter et al. (2011) found that for these two dimensions, only teachers’ enthusiasm for teaching was related to students’ ratings of teachers’ enthusiasm and students’ enjoyment of mathematics. Indeed, teaching enthusiasm is one of the strongest predictors for students’ interest development (Frenzel, Goetz, Lüdtke, P krun, & Sutton, 2009; Lazarides et al., 2018). This connection was mediated by teachers’ enthusiastic behaviour and individual support in the classroom (Keller et al., 2014; Kunter et al., 2013; Lazarides et al., 2018). Taken together, the intrinsic value of teaching that enthusiastic teachers experience seems to particularly contribute to a supportive and warm atmosphere in the classroom. It might be easier for teachers who have more fun while teaching to create a respectful atmosphere and also to put effort into supporting each student individually (Roth, Assor, Kanat-Maymon, & Kaplan, 2007).

The Present Study

Previous studies have predominantly measured teacher effects over longer time periods (e.g., Guo et al., 2012; Hill, Rowan, & Ball, 2013; Kunter et al., 2013). These time periods have the advantage that students' long-term cumulative learning in a subject can be evaluated. However, this approach may also entail some undesired side effects. Most importantly, there are numerous factors that occur across a whole school year that are not under the control of researchers but are likely to influence both independent and dependent variables. In addition, this approach makes it necessary to apply very broad outcome measures, which are again more prone to external influences.

The basic idea of the present investigation was to evaluate the effects of teacher competence and teaching quality in a standardized research design (Authors, 2014a; Authors, 2015a). This investigation is part of a larger research design in which all teachers received the same materials for two predesigned teaching units on similar topics in the area of "floating and sinking" (materials from Hardy, Jonen, Möller, & Stern, 2006), accompanied by a professional development workshop on basic scientific concepts. This made it possible to assess student outcomes (conceptual understanding and interest) closely related to these instructional units. The fact that the topic addressed is demanding for teachers as well as for students makes these instructional units very well-suited for addressing our research questions. Here, the specific teaching unit is the vessel through which teacher performance can be evaluated. This approach has previously been applied in research on teaching quality in mathematics and science education (Hardy et al., 2006; Klieme et al., 2009; Lipowsky et al., 2009). It offers the advantage of involving fewer external (unobserved) factors that could possibly threaten the internal validity of the results.

Research Questions and Hypotheses

With this design, we investigated whether the previous findings described above could be transferred to the field of elementary science education, where the gap between teacher competence and the high demands of everyday teaching practice is more pronounced than in other domains and grade levels (Appleton, 2008; Brobst et al., 2017). We examined the following research questions and hypotheses:

1. What are the effects of teachers' pedagogical content knowledge, constructivist beliefs, self-efficacy, and teaching enthusiasm on students' conceptual understanding of floating and sinking and interest in the teaching units? We expected the aspects of teacher competence to be positively associated with the student outcome measures of conceptual understanding and interest (Hypothesis 1).
2. What are the effects of the four aspects of professional competence on the three basic dimensions of teaching quality? We expected PCK (Hypothesis 2) as well as constructivist beliefs (Hypothesis 3) to be associated with the supportive climate and cognitive activation dimensions. We further expected teaching enthusiasm to be associated with supportive climate (Hypothesis 4), and self-efficacy to be associated with all three teaching quality dimensions (Hypothesis 5).
3. Are the effects of teacher competence on student outcomes mediated by teaching quality? We expected the respective teaching quality dimensions to mediate the relations between teacher competence and student outcomes (Hypothesis 6).

The fundamental idea behind these research questions is that the effects of teacher competence on student outcomes are mediated by what actually happens in the classroom in terms of teaching quality. The research questions build upon previous research conducted within the same research project, which has provided evidence that the teaching quality measures we used are related to student outcomes (Authors, 2014a; 2014b; Authors, 2015b). This work was mainly concerned with methodological questions of how to best assess teaching quality. It also offers indications as to which data source should be used to measure each basic dimension of teaching quality (see Instruments section). Relations between teacher competence and teaching quality as well as between teacher competence and student outcomes, which represent the focus of the present work, have not been examined before.

Method

Design

Participating teachers taught two teaching units on the topic of floating and sinking in their Grade 3 class. This topic is part of the science curriculum in German primary schools. The materials and schedule we used were adapted from an empirically evaluated science curriculum for teaching floating and sinking that was modeled on the principles of inquiry-based science education (Möller & Jonen, 2005). The first unit covered the concept of density, and the second unit focused on the concepts of buoyancy force and displacement. Each unit consisted of 4.5 lessons of 90 minutes that were integrated into regular courses for a duration of about nine weeks (about one 90-minute lesson per week).

The implementation of these units was part of a larger design for evaluating different teaching approaches in science education in German primary schools (Authors, 2015a). As part of this design, the teachers in the study participated in professional development workshops. One part of these workshops (4.5 hours) focused on the scientific concept of density and on how to implement the teaching units and use the materials. This part of the training was the same for all groups. In the second part of the workshops, teachers were randomly assigned to four different conditions. Three of them were given one approach to adaptive teaching, while the fourth served as a control group (Authors, 2015a).

We did not focus on these different approaches in the present study. Instead, we examined the data set as a whole and did not differentiate between treatment conditions. We did not expect the treatment conditions to impact student learning in a mediation model with teaching quality, as the different approaches referred to specific teaching practices that were quite independent of the generic dimensions of teaching quality examined in this study (Authors, 2015b). In all of the four groups, the same inquiry-oriented science curriculum was taught. Thus, all of the treatment conditions were expected to offer a high degree of cognitive activation due to the basic curriculum. There were also no expectations that the other two dimensions of teaching quality would vary systematically between the conditions.

Descriptively, there were only small differences between the treatment groups regarding our major variables of interest. We also conducted empirical tests of whether the treatment conditions differed. Regressing variables of teacher competence or teaching quality on dummy-coded treatment groups did not reveal any significant relationships (using a more conservative alpha level of 10%). We additionally controlled for the various treatment conditions in our analyses to make sure that they did not affect the relations between teacher competence, teaching quality, and student outcomes (see Data Analyses section).

The measures of teacher competence were taken after teachers completed the workshops. One exception was enthusiasm for teaching, which was assessed only prior to the workshops. We included this measure nonetheless because we regard enthusiasm as a rather stable characteristic of teacher motivation (Keller et al., 2014; Kunter et al., 2008) that should not have been affected by the workshops. All measures of teacher competence were collected before teachers started to implement the teaching units. Teaching quality during these units was assessed via student ratings after the first teaching unit (supportive climate and classroom management) and via external observations in one block period of 90 minutes during the first unit (cognitive activation). Pre and post measures of students' conceptual understanding and interest were taken before and after the units.

In summary, this research design included direct measures of teacher competence, which we could link to data on student development in a well-defined content area with pre-post measures and measures of teaching quality during these units from multiple sources.

Participants

The total sample consisted of $N = 54$ teachers and 1,070 students from Grade 3. Each teacher taught one class (one class could not be rated by external observers, see below). About 20 students were assessed per class (min. 10, max. 27). Participating teachers had a mean age of 42.8 years ($SD = 9.2$) and teaching careers that spanned 16.4 years on average ($SD = 8.6$ years). 89% of teachers were female, which reflects the proportion of female teachers in German primary schools. The average student age was 8.8 years ($SD = 0.50$), and 49% of the students were female. The target populations were students and teachers from public primary schools in a German state. Participating schools were located in both urban (61% of classes) and rural areas. Participation in the study was voluntary for both teachers and students. The average participation rate for each classroom was 96%.

Instruments

Teacher competence. We assessed the aspects of teacher competence with questionnaires or standardized tests, respectively.

Pedagogical content knowledge (PCK). We measured teachers' PCK with a standardized test based on the instrument by Lange (2010). It was comprised of five open-ended items from two areas: (1) Knowledge about instructional strategies and representations (sample item: "Please name four different experiments that help to explain the topic 'floating and sinking' to students"), (2) Knowledge about students' subject-specific preconceptions and misconceptions (sample item: "Please name typical misconceptions that students use to explain phenomena of floating and sinking”). According to a coding manual that provided clear and distinct rules and was developed by experts in the field of science didactics, each of the participants’ answers could be categorized as appropriate or not. Appropriate answers were counted and added to a sum score. The internal consistency of the sum score was Cronbach’s α = .67. About one third of participants’ responses were coded twice by two independent raters. Raters reached an average interrater agreement of 83% and an average interrater reliability of ICC = .80. This instrument was shown to be sensitive to the effects of a professional development workshop for teachers, which provided evidence for the validity of this instrument (Authors, 2015c).

Constructivist beliefs. Teachers’ constructivist beliefs were measured with a four-item scale that was based on scales by Warwas et al. (2011) and Staub and Stern (2002). The items focused on the degree to which teachers believed in the advantages of independent and discursive learning processes (e.g., “Children learn especially well when they are allowed to develop their own ideas and go their own way while learning”; Cronbach’s α = .72). Items on enthusiasm and beliefs were rated on a 4-point Likert scale ranging from 1 (*strongly disagree*) to 4 (*strongly agree*).

Self-efficacy. Teacher self-efficacy was measured with nine items that were modified from a scale by Schmitz and Schwarzer (2000) that is well-established and widely used in German speaking countries (sample item: “How confident do you feel teaching all relevant subject content to even the most difficult students?”; Cronbach’s α = .80). It captures teachers’ confidence in managing a variety of situations in everyday school life with a focus on classroom instruction. The metric for this scale ranged from 0 to 100% agreement with the items. Previous studies provided evidence for the validity of this scale. It is negatively correlated with burnout (Maslach Burnout Inventory), teachers self-reported work strain, and the time teachers spend with their students in extracurricular activities (Schmitz & Schwarzer, 2000). The teacher self-efficacy scale outperformed a general self-efficacy scale regarding the prediction of these measures (Schmitz & Schwarzer, 2000). It has also been shown to predict teachers' actual behavior in the classroom (teaching quality) (Holzberger et al., 2013; Praetorius et al., 2017).

Teaching enthusiasm. We measured teaching enthusiasm with a six-item scale by Kunter (2008; e.g., "Teaching is a great pleasure for me."; Cronbach's α = .81). The studies by Kunter et al. (2011) and Lazarides, Gaspard & Dicke (2018) provided evidence for the factorial and predictive validity of the scale.

Teaching quality. Previous research has revealed that different sources of data are necessary to properly assess the different dimensions of teaching quality in elementary school: Results from multilevel factor analyses have shown that third graders provide sufficiently differentiated ratings of teaching quality in survey-based assessments (Authors, 2014b). However, it has been shown that third graders have problems with rating the potential for cognitive activation as such judgements might exceed their understanding of pedagogy (Authors, 2014a). We thus decided to use external (video-)observer ratings to assess the potential for cognitive activation and student ratings to assess supportive climate and classroom management.

We measured *supportive climate* with nine items that were related to positive teacher-student interactions and covered teachers' warmth and friendliness, encouragement, and constructive feedback (sample item: "Our science teacher tells me how to do better when I make a mistake"). Items were evaluated regarding their internal consistency (Cronbach's α = .90), the proportion of variance attributable to the classroom level of analysis (ICC = .16), and the reliability of the classroom aggregate regarding students' agreement within classes (ICC2 = .78, according to formula by Lüdtke et al., 2009, with values > .70 usually regarded as indicating acceptable reliability, LeBreton & Senter, 2008). The *classroom management* scale included five items (Cronbach's α = .91, ICC = .25, ICC2 = .86) related to the absence of disciplinary problems and disruptions during classroom instruction (sample item: “In our science class students are quiet when the teacher speaks”).

We measured the potential for *cognitive activation* with high-inference ratings of external observers on a rating scale called “challenging tasks and questions.” This scale was designed to measure how the teacher actually succeeded in creating a learning environment that challenges students to engage in reasoning about the concepts taught through science practice and how the teacher uses experiments and questions to initiate conceptual change in students. More specifically, the scale’s indicators cover the activation of students’ prior knowledge and concepts, and the extent to which the teacher posed open questions that challenge students to more deeply cognitively engage with the learning content. Additionally, observers had to rate whether students were given the chance to develop their own ideas to explain the findings of experiments, whether students were encouraged to explain contradictory observations, and whether single observations were linked to the respective concept. Similar indicators have been shown to be crucial in secondary math instruction (Lipowsky et al., 2009) and were adapted to capture the potential for cognitive activation in the context of elementary science instruction (Authors, 2014a).

Raters received extensive training (approximately 40 hrs) and assigned their ratings on one 4-point Likert scale ranging from 1 (*strongly disagree*) to 4 (*strongly agree*), integrating the indicators described above in accordance with a coding manual. Interrater reliability was sufficient (ICC = .77 for two independent raters; Shrout & Fleiss, 1979). Observations were made either from video recordings ($n = 37$) or live in the classroom ($n = 16$) for teachers who did not agree to be recorded (see Authors, 2014a).

Previous studies on the psychometric properties of the instruments have provided evidence for the validity of the three teaching-quality scales regarding the discriminant validity of student ratings, convergence between different data sources, and the prediction of student outcomes (Authors, 2014a, 2014b).

Student outcomes and covariates. Students' interest and conceptual understanding of the content covered in the teaching units were assessed with pre-post measures.

To measure students' *prior interest* in science education, we used a four-item scale (e.g., "I put effort into science class because it is fun"; Cronbach's α = .89, ICC = .20) that was based on a scale by Blumberg (2008). Student interest after the science classes was measured using the same items with a different stem focusing on students' *interest in the teaching unit* (e.g., "I put effort into the topic of floating and sinking because it was fun"; Cronbach's α = .91, ICC = .16). Students' *conceptual understanding* of floating and sinking was assessed with standardized tests. Test items were adapted from existing instruments by Hardy et al. (2006) and Kleickmann et al. (2010). The pretest comprised 16 items (EAP/PV reliability = .52, ICC = .06), and the posttest comprised 13 items (EAP/PV reliability = .76, ICC = .19). EAP/PV reliability was computed using a formula from Adams (2005) and can be interpreted like Cronbach's alpha, with values above .70 generally regarded as acceptable (Kline, 2000). Because the pretest on conceptual understanding exhibited insufficient reliability, we additionally tested students' science competence and cognitive abilities and used them as covariates to control for pre-existing differences between students (see below). Regarding validity, these items have been shown to be sensitive to instruction (Naumann, Hochweber, & Hartig, 2014). In addition, experts from educational practice and research in science education have judged the items to be valid and highly relevant to the topic of floating and sinking. Items were scored dichotomously or polytomously, and the two tests were scaled separately by applying the partial credit model each time. Student parameters were estimated with weighted likelihood estimates (Warm, 1989).

We assessed students' prior *science competence* with an adapted version of the TIMSS test (Martin, Mullis, & Foy, 2008) that fit the 1PL Rasch model (13 items; EAP/PV reliability = .70; ICC = .14). *Cognitive abilities* were assessed with the CFT 20-R (56 items, Cronbach's α = .72; ICC = .10; Weiß, 2006), a German version of the culture fair intelligence tests.

Data Analyses

Due to the clustered data structure, with individual students nested within classes at Level 1 and teacher competence and teaching quality measures at Level 2, we used multilevel analyses. We specified doubly manifest models according to the framework proposed by Lüdtke, Marsh, Robitzsch, and Trautwein (2011), as this was supposed to be the best solution, considering the relatively small sample size at Level 2 and the almost full coverage of units within classes at Level 1 (96% on average). Predictions of student outcomes were estimated as random-intercept models, as we did not expect the slopes of Level 1 covariates to vary across cluster units (Raudenbush & Bryk, 2010). All regression models were estimated in Mplus 7 (Muthén & Muthén, 1998–2012) using robust maximum likelihood estimation (MLR; Yuan & Bentler, 2000). Coefficients of determination were estimated as the ratio between explained and total variance separately for each level of analysis.

To answer Research Question 1, we estimated two sets of multilevel regression models: one for students' conceptual understanding posttest scores as the dependent variable and one with students' subject-related interest in the two teaching units as the dependent variable. Student level covariates were introduced as group-mean-centered Level-1 predictors and additionally as classroom-aggregated predictors at Level 2. We also introduced dummy-coded treatment conditions (see Design section) at Level 2 to control for potential differences between groups.

To answer Research Questions 2 and 3, we further evaluated the models from Research Question 1. In this model, we regressed the basic dimensions of teaching quality on the teacher competence variables (Hypotheses 2 to 5) and we introduced the basic dimensions as predictors of student outcomes (see Figure 1). Thus, we specified a 2-2-1 mediation design according to Preacher et al.'s (2010) framework (Hypothesis 6). We tested the indirect effects of teacher competence on student outcomes via teaching quality using the *model indirect* command implemented in Mplus (Muthén & Muthén, 1998–2012), which uses the delta method (MacKinnon, 2008, p. 92).

The issue of missing values requires careful consideration (Enders, 2010). In our study, a relatively small amount of missing data occurred at the level of individual students (M = 7.97%, Range = 6.8–9.7%, see Table 1). Missing data on teacher variables ranged between two and three cases (see Table 1). Missing values were generated when students or teachers did not attend school on the day the measurements were taken. Absence was mostly due to illness or, for student measures, because students changed classes or schools. One missing value in the external classroom observations occurred for organizational reasons. There was no missing data for classroom-level aggregates of individual student data. There was no indication of a systematic accumulation of missing data patterns across scales or measurement points. We used a full information maximum likelihood algorithm (FIML; Arbuckle, 1996) to deal with missing data in all regression models.

Results

Descriptive statistics are presented in Table 1 for all variables. A first look at the bivariate correlations between the different aspects of teacher competence reveals rather small to zero correlations between these variables (Table 2). There was only one significant correlation between self-efficacy and teaching enthusiasm. The same applies to the dimensions of teaching quality. Only the correlation between supportive climate and classroom management (both measured via student ratings) was significant. Thus, teachers differ with regard to the different aspects of professional competence. For example, a teacher with a high PCK is not necessarily one who is also very enthusiastic about teaching. For the bivariate relations in Table 2, there were also some small to medium correlations between the four aspects of teacher competence and student outcomes.

Effects of Teacher Competence

To evaluate the effects of teacher competence on student achievement (i.e., conceptual understanding) and interest (Research Question 1), we first introduced each aspect of teacher competence in the regression analyses in a stepwise fashion (Table 3; Models 1a-d and Models 2a-d). Teachers' self-efficacy was a significant predictor of student achievement and student interest (Models 1c and 2c). PCK and teacher enthusiasm were related to student interest but not to achievement (Models 2a and 2c). Constructivist beliefs were not related to either of the two outcomes (Models 1d and 2d). Thus, these results are only partly in line with Hypothesis 1. In Models 3 and 4, we evaluated the unique contribution of each predictor over and above all of the other predictors. The effects did not change substantially in comparison with the single predictor models. Thus, the previously significant predictors also made a unique contribution to explaining the student outcomes.

The Role of Teaching Quality

To answer Research Question 2, we examined relations between teacher competence and teaching quality. The results of the full multilevel mediation model (Table 4 and Figure 1) revealed significant effects of teachers' pedagogical content knowledge on supportive climate but not on cognitive activation. Thus, the results are only partly in line with Hypothesis 2. Contrary to our expectations in Hypothesis 3, we did not find significant relations between constructivist beliefs and teaching quality. Teaching enthusiasm was significantly related to the supportive climate dimension and—as expected in Hypotheses 4 and 5—teacher self-efficacy predicted all three basic dimensions. In the full mediation model, all predictors were introduced simultaneously. Thus, significant predictors made a unique contribution to teaching quality over and above the others.

Figure 1 presents a review of the mediation analyses (Research Question 3). Detailed results in Table 4 show that the effect of self-efficacy on student achievement decreased and was no longer significant when teaching quality was controlled for. The same was true for the effects of PCK, self-efficacy and teaching enthusiasm on student interest. The evaluation of indirect effects revealed that the main effect of self-efficacy on student achievement was mediated by teaching quality ($\beta_{indirect} = .25$, $p < .05$), whereas the indirect effect of self-efficacy on student interest was not significant ($\beta_{indirect} = .08$, $p > .05$). The indirect effect of PCK on student interest via teaching quality was significant ($\beta_{indirect} = .12$, $p < .05$). The effect of teaching enthusiasm on student interest was mediated by teaching quality as well ($\beta_{indirect} = .14$, $p < .05$). These results are in line with Hypothesis 6.

Discussion

In this study, we examined relations between teacher competence, teaching quality, and student outcomes in elementary science education. The first aim of this article was to identify specific, measurable aspects of teacher competence that help teachers succeed in promoting student achievement and interest in complex settings related to science learning. We further expected teaching quality to play an important role in this link: Competent teachers are able to provide high-quality instruction, which in turn affects student progress. Our study was based on several theoretical concepts and models from research on secondary math and science instruction. We were able to successfully apply these concepts to elementary science education, thus adding knowledge to a very important field of research. In contrast to most secondary school teachers, elementary school teachers are often generalists who do not necessarily have an academic background in science-related subjects, posing particular challenges in the field of elementary science education (Appleton, 2008; Johnston & Ahtee, 2006; Krämer et al., 2015; Rice, 2005). Complex mechanisms underlie younger students' understanding of science phenomena. Teachers have to not only understand these phenomena themselves (which is challenging enough) but also have to be willing and able to explain them to their students. The relations between different aspects of teacher competence and classroom processes might therefore look different in elementary science classes. Teacher's motivation (enthusiasm and self-efficacy) might play a role that is just as important as the more cognitive aspects of teacher competence (like knowledge and beliefs) on which much of the research from secondary schools has focused so far. However, our results show that similar conceptualizations of teacher competence and teaching quality apply to different domains and grade levels. Before we discuss the results in detail, we will highlight some overarching findings.

First, we were indeed able to confirm that several aspects of teacher competence are related to student outcomes. These effects were at least partly mediated by the “basic dimensions” as generic aspects of teaching quality (Klieme et al., 2009). Thus, the positive effects of teacher competence can to some extent be explained by what successful teachers actually do in the classroom in terms of teacher-student interactions during various tasks and instructional settings. Additionally, it seems that variables more proximal to student learning (teaching quality) have greater explanatory power for student outcomes than the rather distal aspects of teacher competence.

Second, we primarily found specific, unique effects. For example, each of the aspects of PCK, self-efficacy, and teaching enthusiasm made a unique contribution to explaining supportive climate in the classroom. Professional competence does not represent a single global teacher characteristic (Kunter et al., 2013). Instead, several aspects of teacher competence need to be considered when explaining teaching quality and student progress. Teachers differ with regard to their individual strengths and weaknesses. To our knowledge, this is the first study to examine the specific effects of a comprehensive set of teacher and teaching characteristics in the field of elementary science education.

Pedagogical Content Knowledge: Not Related to Achievement or Cognitive Activation

PCK was identified as a relevant predictor of teachers’ cognitive activation and supportive climate in the classroom (Baumert et al., 2010; Förtsch et al., 2017) in secondary classes. Although science specific PCK might be limited among elementary school teachers, it was also found to predict students’ conceptual understanding and subject-related interest in elementary science education (Lange et al., 2012). In our study, PCK was related only to supportive climate and student interest. How can these different results be explained? We discuss two potential explanations, both of which are related to the standardized design of the present study.

First, we assessed teachers' PCK in a very distinct content area (knowledge on teaching floating and sinking). The fact that all teachers had previously dealt with how to implement this teaching unit in the classroom might have led to restricted variance among teachers' PCK on this topic. This would lead to smaller correlations with other constructs (in addition to the rather low reliability of the PCK measure we used, which has the same effect). However, this does not explain why effects were found for supportive climate and student interest, which brings us to another explanation. The PCK construct we measured focused on two main areas: (1) teachers' knowledge of how the topic of floating and sinking can best be explained to students and (2) teachers' knowledge about students' preconceptions and misconceptions in this content area. The teaching units that teachers had to implement included specific cognitively activating experiments, tasks, and worksheets for the students (Hardy et al., 2006). While these curriculum materials do certainly not determine a teacher's instruction in the classroom, they might have led to a limited variation in how teachers dealt with the topic in their classes. Thus, the first component of teachers' PCK—the part that enables them to choose appropriate tasks and experiments when preparing for upcoming lessons—was probably less important as a source of variance between teachers in the present study. On the other hand, teacher-student discourse and one-on-one scaffolding did occur during the teaching units. Thus, the second part of PCK became relevant during teacher-student interactions when students who held certain preconceptions and misconceptions asked for help and the teacher had to provide individual support (van de Pol, Volman, & Beishuizen, 2010). This was reflected in the association between teachers' PCK scores and students' ratings of supportive climate. A higher degree of supportive climate was in turn positively associated with students' interest. The path from PCK through supportive climate to student interest has also been found in previous studies on mathematics classes (Kunter et al., 2013).

Unfortunately, we could not test these interpretations because we were not able to split the short PCK measure into two scales reflecting the aforementioned components. Future research will have to further explore the complex relations between the subfacets of PCK and teaching quality.

Constructivist Beliefs: Unrelated to Student Outcomes and Teaching Quality

Teacher beliefs about teaching and learning play an important role in research on mathematics and science teaching (Handal, 2003; Jones & Carter, 2007). However, results on the relations between teachers' beliefs and what teachers actually do in the classroom are mixed. Differently from PCK, the non-significant effects we found for beliefs cannot be attributed to our study design, as teacher beliefs should not depend on whether the teaching units were predesigned or not. Additionally, the descriptive results show that constructivist beliefs were not particularly low in our sample.

Theoretical considerations suggest that constructivist beliefs can be beneficial for student learning if they are accompanied by an increased awareness of the importance of students' misconceptions (Vosniadou, 2013). This could in turn lead to teaching behavior that explicitly addresses students' misconceptions, which would be reflected in increased individual support in classrooms of teachers with more constructivist beliefs. However, our results show that holding certain beliefs does not necessarily have an impact on the actual behavior in the classroom in terms of teaching quality.

Earlier studies have challenged the importance of constructivist teacher beliefs for teaching quality and student learning too: Kunter et al. (2013) found no specific effect of constructivist beliefs on student achievement and only a negative correlation with classroom management when other facets of teacher competence were controlled for. In a study by Kleickmann et al. (2016), the effects of professional development workshops on student achievement were not mediated by teachers' beliefs but by teachers' instructional quality in the classroom. The authors suggested that having certain beliefs about the nature of science teaching and learning might not be sufficient to achieve high instructional quality. Our results are in line with this interpretation.

Teacher Self-Efficacy: The Most Important Predictor of Student Outcomes and Teaching Quality

Self-efficacy was the most prominent predictor of student development in our study. It was the only aspect of teacher competence that was connected to the development of students' conceptual understanding. In addition, self-efficacy was the only competence measure that was related to external observer ratings of cognitive activation.

Previous research has underscored the role of teachers' self-efficacy for teacher competence (Holzberger et al., 2013, 2014). However, the effects of self-efficacy in previous studies were not as pronounced as previously expected (Klassen et al., 2011; Klassen & Tze, 2014). The fact that we found substantial associations with student outcomes and teaching quality might have to do with the subject matter and grade level that we focused on. Self-efficacy always refers to people's beliefs about whether they are able to deal with certain situations. The more challenging a situation, the more relevant self-efficacy becomes. Teaching science is a challenging task for many elementary school teachers (Davis et al., 2006). It is plausible that the teachers who performed well were the ones who felt confident in their ability to manage the challenge of implementing a demanding teaching unit in science education. It was nonetheless remarkable that a rather general measure of self-efficacy predicted student learning over and above the other measures of teacher competence. However, our results are also in line with Zee and Koomen's (2016) suggestion that classroom processes might mediate the effects of teacher self-efficacy. For future research, it would be desirable to apply more specific measures of self-efficacy (e.g., efficacy for specific instructional strategies), which would also allow more specific predictions of the different dimensions of teaching quality (Tschannen-Moran & Hoy, 2001). Similarly, given the special role of elementary science in relation to other subjects, it would be desirable to apply measures with a more specific focus on efficacy for science education in future studies.

Teaching Enthusiasm: Teachers' Motivation Fosters Students' Motivation

Teaching enthusiasm refers to intrinsic aspects of teachers' motivation, which has previously been linked to the development of students' interest (Keller et al., 2014). This effect has also been referred to as "emotional transmission in the classroom" (Frenzel et al., 2009). Teachers who experience enjoyment and pleasure during teaching are better able to raise students' subject-related interest. Our study lends further support to these findings. First, we were able to confirm that this effect can be generalized to the field of elementary science education. Second, the effect arose in the rather brief amount of time it took to teach two units on one topic. Third, our study went beyond merely showing that this effect is mediated by teachers' expressed enthusiasm in the classroom (Keller et al., 2014; Frenzel et al., 2009) or autonomy-supportive teaching (Roth et al., 2007). We were further able to show that the basic dimension of supportive climate as a broader and more general aspect of teaching quality also contributed to this effect (cf. Kunter et al., 2013). Like in the case of teacher self-efficacy, it would be interesting to apply measures that are more specific to the case of elementary science education in future studies. However, we also believe that the present findings offer a very good starting point for more domain-specific analyses by showing that the more general aspects of teacher enthusiasm and self-efficacy we considered in our study indeed matter for teaching quality and student outcomes in elementary science.

Value of the Study for Research on Teaching Quality: Antecedents and Consequences

Our study illustrates that it is fruitful to consider different sources of data for different aspects of teaching quality. We included different sources of data and were able to confirm the expected specific and unique effects of the three basic dimensions of cognitive activation, supportive climate, and classroom management (Authors, 2015b; Authors, 2014b). These results are in line with other studies involving various school subjects and grade levels (Kunter et al., 2013; Pianta & Hamre, 2009). In addition, we added to knowledge about the specific antecedents of high-quality teaching in elementary science education in terms of teachers' professional competence. Identifying such competencies is particularly important for promoting teaching quality (see Implications section).

It is in line with our expectations that the teacher variables that were related to student interest (PCK, self-efficacy, and teaching enthusiasm) were also related to supportive climate. On the other hand, the teacher variable connected to student achievement (self-efficacy) was also related to classroom management and cognitive activation. This pattern strengthens our interpretation that there are systematic relations between teacher competence, teaching quality, and student outcomes and provides further evidence for the validity of our results.

Further Strengths and Limitations of the Study

We were able to link direct measures of several facets of teacher competence to different aspects of teaching quality and two important student outcomes in a longitudinal design. Effects were estimated in a standardized research design, which increased the internal validity of our study. Teacher competence develops over the course of the school year, and the impact of teaching experience on teacher competence can be as pronounced as the impact of teacher competence on teaching quality (Holzberger et al., 2013). Observing teacher effects in a predesigned teaching unit controls for these bidirectional effects to a large degree purely by design. Nevertheless, our study remains correlational in nature. We have good reason to believe that the effects we detected move in the causal direction suggested above, but our research design did not allow us to test this. We propose that teacher competence influences teaching quality and, in turn, student outcomes, but this interpretation leaves out potential bidirectional effects. Research has shown, for example, that teachers' self-efficacy is affected by their previous experiences in the classroom (Holzberger et al., 2013; Zee & Koomen, 2016). The same might apply to student outcomes: We can imagine that teachers of students who perform well (for any reason) will feel more effective. In this interpretation, self-efficacy would not be the cause but rather the outcome of student achievement. Our research design (specific standardized teaching unit, strong control variables) makes this interpretation less likely, but it cannot be definitively ruled out. Future research is needed to further examine the effects described in the present paper.

The advantages of our research design also go along with some drawbacks. We put a lot of effort into ruling out the possibility that the different treatment conditions within the larger research design had an effect on our results. However, a completely untreated sample would be desirable for future studies in this area. Additionally, as in many other studies, participation in our study was voluntary for teachers. We cannot rule out that teachers in our sample were more competent than those who did not participate, which would also limit the generalizability of our findings.

Implications for Practice and Future Research

Motivational constructs, namely teacher motivation (in the form of self-efficacy and enthusiasm) and student motivation (in the form of subject-related interest), played an important role in the present investigation. Researchers have underscored the important role of motivation in schools for a long time (Richardson, Watt, & Karabenick, 2014). Policymakers have called for more highly educated students to fill STEM-related jobs, which are considered crucial for the development of modern societies. Thus, several countries are struggling with the question of how to increase students' interest in science-related topics. Research shows that strong results in science competence on the country level do not necessarily go along with high interest in science among students (Martin, Mullis, Foy, & Hooper, 2016; OECD, 2016). In addition, in the field of science education, a steady decline in student interest has been observed with increasing grade levels. On a policy level, focusing more on teacher competence seems to be a promising way to deal with these issues. Competent, highly motivated teachers might be able to help prevent this decline, and teacher education and professional development seem to be good levers for promoting students' science-related interest.

An important feature of our research design is that it allowed us to identify important practical implications because we considered the development of conceptual understanding and interest in a distinct content area rather than in a fairly abstract competence area. This is something that teachers have to deal with in their everyday teaching practice. Our approach of directly measuring teacher competence and teaching quality offers important advantages over other methods of estimating teacher quality, such as value added measures (Polikoff, 2015). Most importantly, teachers differ in their individual strengths and weaknesses, which was reflected by the low correlations between the different aspects of teacher competence. The specific assessment of these aspects offers valuable starting points for professional development. Thus, at the level of everyday school practice, the promotion of these specific aspects of teacher competence seems promising with a view to improving the quality of classroom instruction. The aspects of teacher competence that were effective in this study are not static, unchangeable teacher characteristics. Rather, previous research has provided rich evidence that each of the aspects described above can be developed and fostered in professional development programs (e.g., Borko, 2004; Kleickmann et al., 2016; Timperley, 2008).

More specifically, the approaches we used to measure teaching quality (observations and student surveys) offer rich information that can be used to provide teachers with detailed feedback (Borko, Jacobs, Eiteljorg, & Pittman, 2006). This is particularly important, as feedback is known to be one of the most powerful tools for professional development and improving teacher competence. For example, teacher training that uses such measures of teaching quality will be very well able to identify teachers' own strengths and weaknesses in everyday classroom instruction. This is a major advantage of the present approach in comparison to other indicators of teacher quality such as value-added measures. However, this assumption—that improving teacher competence can increase teaching quality and student outcomes—will have to be examined in greater detail in further, preferably experimental, studies.

Conclusion

In the present study, we examined the effects of teacher competence on teaching quality and student outcomes. The most prominent aspect of teacher competence was teachers' general self-efficacy, which predicted students' conceptual understanding of taught content as well as their subject-related interest. Teachers' pedagogical content knowledge and enthusiasm for teaching were also related to student interest, but not to achievement. Moreover, these effects were mediated by the three basic dimensions of teaching quality: cognitive activation, supportive climate, and classroom management. These results help us better understand the mechanisms behind the effects of teachers' personal characteristics. They can also be informative for future efforts to improve educational quality.

References

Anderson, L. M., Blumenfeld, P. C., Pintrich, P. R., Clark, C. M., Marx, R. W., & Peterson, P. L. (1995). Educational psychology for teachers: Reforming our courses, rethinking our roles. *Educational Psychologist, 30*, 143–157.

Appleton, K. (2008). Elementary science teaching. In S. K. Abell & N. G. Lederman (Eds.), *Handbook of research on science education* (pp. 493–535). New York, NY: Routledge.

Arbuckle, J. L. (1996). Full information estimation in the presence of incomplete data. In G. A. Marcoulides, & R. E. Schumacker (Eds.), *Advanced structural equation modeling* (pp. 243–277). Mahwah, NJ: Lawrence Erlbaum.

Authors (2014a) [details removed for peer review].

Authors (2014b) [details removed for peer review].

Authors (2015a) [details removed for peer review].

Authors (2015b) [details removed for peer review].

Authors (2015c) [details removed for peer review].

Bandura, A. (1995). *Self-efficacy in changing societies*. Cambridge University Press.

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., et al. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal, 47*, 133–180.

Baumert, J., & Kunter, M. (2013). The COACTIV model of teachers' professional competence. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, & M. Neubrand (Eds.), *Cognitive activation in the mathematics classroom and professional competence of teachers* (pp. 25–48). New York: Springer.

Blumberg, E. (2008). *Multikriteriale Zielerreichung im naturwissenschaftsbezogenen Sachunterricht der Grundschule* [Multi-criterial goal attainment in science education in primary school] (Doctoral dissertation). University of Münster, Germany.

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher, 33*, 3–15.

Borko, H., Jacobs, J., Eiteljorg, E., & Pittman, M. E. (2008). Video as a tool for fostering productive discussions in mathematics professional development. *Teaching and Teacher Education, 24*, 417–436.

Bromme, R. (2001). Teacher expertise. In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social and behavioral sciences* (pp. 15459–15465). Oxford: Elsevier.

Brobst, J., Markworth, K., Tasker, T., & Ohana, C. (2017). Comparing the preparedness, content knowledge, and instructional quality of elementary science specialists and self- contained teachers. *Journal of Research in Science Teaching, 54*, 1302–1321.

Brophy, J. (2000). *Teaching*. Brussels: International Academy of Education.

Brophy, J., & Good, T. L. (1986). Teacher behavior and student achievement. In M. Wittrock (Ed.), *Handbook of research on teaching* (pp. 340–370). New York: Macmillan Library.

Davis, E. A., Petish, D., & Smithey, J. (2006). Challenges new science teachers face. *Review of Educational Research, 76*, 607–651.

Davis, H. A. (2003). Conceptualizing the role and influence of student-teacher relationships on children's social and cognitive development. *Educational Psychologist, 38*, 207–234.

Deci, E. L., & Ryan, R. M. (2000). The ''what'' and ''why'' of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry, 11*, 227–268.

Depaepe, F., Verschaffel, L., & Kelchtermans, G. (2013). Pedagogical content knowledge: A systematic review of the way in which the concept has pervaded mathematics educational research. *Teaching and Teacher Education, 34*, 12–25.

Dubberke, T., Kunter, M., McElvany, N., Brunner, M., & Baumert, J. (2008). Lerntheoretische Überzeugungen von Mathematiklehrkräften [Mathematics teachers' beliefs about learning]. *Zeitschrift für Pädagogische Psychologie, 22*(3), 193–206.

Duschl, R. A. & Bybee, R. W. (2014). Planning and carrying out investigations: An entry to learning and to teacher professional development around NGSS science and engineering practices. *International Journal of STEM Education*, 1:12. Retrieved from https://link.springer.com/article/10.1186/s40594-014-0012-6.

Emmer, E. T., & Stough, L. M. (2001). Classroom management: a critical part of educational psychology, with implications for teacher education. *Educational Psychologist, 36*, 103–112.

Enders, C. K. (2010). *Applied missing data analysis*. New York: Guilford.

Fives, H. & Gill, M. G. (2015). *International handbook of research on teachers´ beliefs*. New York: Taylor & Francis.

Fives, H., & Buehl, M. M. (2012). Spring cleaning for the "messy" construct of teachers' beliefs: What are they? Which have been examined? What can they tell us? In K. R. Harris, S. Graham, & T. Urdan (Eds.), *APA educational psychology handbook: Individual differences and cultural and contextual factors* (pp. 471–495). Washington: American Psychological Association.

Förtsch, C., Werner, S., von Kotzebue, L., & Neuhaus, B. J. (2016). Effects of biology teachers' professional knowledge and cognitive activation on students' achievement. *International Journal of Science Education, 38*, 2642–2666.

Frenzel, A. C., Goetz, T., Lüdtk , O., Pekrun, R., & Sutton, R. (2009). Emotional transmission in the classroom: Exploring the relationship between teacher and student enjoyment. *Journal of Educational Psychology, 101*, 705–716.

Gitomer, D. & Zisk, R. C. (2015). Knowing what teachers know. *Review of Research in Education, 39*, 1–53.

Gomez-Zwiep, S. (2008). Elementary teachers' understanding of students' science misconceptions: Implications for practice and teacher education. *Journal of Science Teacher Education*, *19*, 437-454.

Greenfield, D. B., Jirout, J., Dominguez, X., Greenberg, A., Maier, M., & Fuccillo, J. (2009). Science in the preschool classroom: A programmatic research agenda to improve science readiness. *Early Education & Development, 20*, 238-264.

Guo, Y., Connor, C. M., Yang, Y., Roehrig, A. D., & Morrison, F.J. (2012). The effects of teacher qualification, teacher self-efficacy, and classroom practices on fifth graders' literacy outcomes. *The Elementary School Journal, 113*, 3–24.

Guo, Y., Dynia, J. M., Yeager Pelatti, C., & Justice, L. M. (2014). Self-efficacy of early childhood special education teachers: Links to classroom quality and children's learning for children with language impairment. *Teaching and Teacher Education, 39*, 12–21.

Handal, B. (2003). Teachers' mathematical beliefs: A review. *The Mathematics Educator, 13*(2), 47–57.

Hardy, I., Jonen, A., Möl er, K. & Stern, E. (2006). Effects of instructional support within constructivist learning environments for elementary school students' understanding of "floating and sinking". *Journal of Educational Psychology, 98*, 307–326.

Hattie, J. (2009). *Visible learning*. New York: Routledge.

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

Holzberger, D., Philipp, A., & Kunter, M. (2013). How teachers' self-efficacy is related to instructional quality: A longitudinal analysis. *Journal of Educational Psychology, 105*, 774–786.

Holzberger, D., Philipp, A., & Kunter, M. (2014). Predicting teachers' instructional behaviors: The interplay between self-efficacy and intrinsic needs. *Contemporary Educational Psychology, 38*, 100-111.

Johnston, J., & Ahtee, M. (2006). Comparing primary student teachers' attitudes, subject knowledge and pedagogical content knowledge needs in a physics activity. *Teaching and Teacher Education, 22*, 503–512.

Jones, M. G., & Carter, G. (2007). Science teacher attitudes and beliefs. In S. K. Abell & N. G. Lederman (Eds.), *Handbook of research on science education* (pp. 1067–1104). Mahwah: Lawrence Erlbaum Associates.

Keller, M. K., Goetz, T., Becker, E. S., Morger, V., & Hensley, L. (2014). Feeling and showing: A new conceptualization of dispositional teacher enthusiasm and its relation to students' interest. *Learning and Instruction, 33*, 29–38.

Keller, M., Neumann, K., & Fischer, H. (2016). The impact of physics teachers' pedagogical content knowledge and motivation on students' achievement and interest. *Journal of Research in Science Teaching, 54*, 1–29.

Klassen, R. M. & Tze, V. M. C. (2014). Teachers' self-efficacy, personality, and teaching effectiveness: A meta-analysis. *Educational Research Review, 12*, 59–76.

Klassen, R. M., Tze, V. M. C., Betts, S. M., & Gordon, K. A. (2011). Teacher efficacy research 1998–2009: Signs of progress of unfulfilled promise? *Educational Psychology Review, 23*, 21–43.

Kleickmann, T., Richter, D., Kunter, M., Elsner, J., Besser, M., Krauss, S., & Baumert, J. (2013). Teachers' content knowledge and pedagogical content knowledge: The role of structural differences in teacher education. *Journal of Teacher Education, 64*, 90–106.

Kleickmann, T., Tröbs , S., Jonen, A., Vehmeyer, J. & Mö̈ler, K. (2016). The effects of expert scaffolding in elementary science professional development on teachers' beliefs and motivations, instructional practices, and student achievement. *Journal of Educational Psychology, 108*, 21–42.

Kleickmann, T., Vehmeyer, J., & Mö̈ler, K. (2010). Zusammenhänge zwischen Lehrervorstellungen und kognitivem Strukturieren im Unterricht [Relations between teacher conceptions and features of scaffolding]. *Unterrichtswissenschaft, 38*, 210–228.

Klieme, E., Pauli, C., & Reusser, K. (2009). The Pythagoras study: investigating effects of teaching and learning in Swiss and German mathematics classrooms. In T. Janik & T. Seidel (Eds.), *The power of video studies in investigating teaching and learning in the classroom* (pp. 137–160). Münster, Germany: Waxmann.

Kline, P. (2000). *Handbook of psychological testing*. New York/London: Routledge.

Kounin, J. (1970). *Discipline and group management in classrooms*. New York: Holt, Rinehart, & Winston.

Krauss S., Brunner M., Kunter M., Baumert J., Blum W., Neubrand M., Jordan A. (2008). Pedagogical content knowledge and content knowledge of secondary mathematics teachers. *Journal of Educational Psychology, 100*, 716–725.

Krämer, P., Nessler, S. H. & Schlüter, K. (2015). Teacher students' dilemmas when teaching science through inquiry. *Research in Science & Technological Education, 33*, 325–343.

Kunter, M., Baumert, J. & Kölle , O. (2007). Effective classroom management and the development of subject-related interest. *Learning and Instruction, 17*, 494–509.

Kunter, M., Frenzel, A., Nagy, G., Baumert, J., & Pekrun, R. (2011). Teacher enthusiasm: Dimensionality and context specificity. *Contemporary Educational Psychology, 36*, 289–301.

Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T., & Hachfeld, A. (2013). Professional competence of teacher: Effects on instructional quality and student development. *Journal of Educational Psychology, 105*, 805–820.

Kunter, M., Tsai, Y.-M., Klusmann, U., Brunner, M., Krauss, S., & Baumert, J. (2008). Students' and mathematics teachers' perceptions of teacher enthusiasm and instruction. *Learning and Instruction, 18*, 468–482.

Kyriakides, L., Christoforou, C., & Charalambous, C. Y. (2013). What matters for student learning outcomes: A meta-analysis of studies exploring factors of effective teaching. *Teaching and Teacher Education, 36*, 143–152.

Lange, K. (2010). *Zusammenhänge zwischen naturwissenschaftsbezogenem fachspezifisch-pädagogischem Wissen von Grundschullehrkräften und Fortschritten im Verständnis naturwissenschaftlicher Konzepte bei Grundschülerinnen und -schülern* [Relations between science related pedagogical content knowledge of elementary school teachers and students conceptual understanding] (Doctoral dissertation). University of Münster, Germany.

Lange, K., Kleickmann, T., Tröbs , S., & Mö ler, K. (2012). Fachdidaktisches Wissen von Lehrkräften und multiple Ziele im naturwissenschaftlichen Sachunterricht [Teachers' pedagogical content knowledge and academic outcomes in science education]. *Zeitschrift für Erz ehungswissenschaft, 15*, 55–75.

Lazarides, R., Gaspard, H. & Dicke, A.-L. (2018). Dynamics of classroom motivation: Teacher enthusiasm and the development of math interest and teacher support. *Learning and Instruction, 60*, 126–137.

LeBreton, J. M., & Senter, J. L. (2008). Answers to twenty questions about interrater reliability and interrater agreement. *Organizational Research Methods, 11*, 815–852.

Leuchter, M., Saalbach, H., & Hardy, I. (2014). Designing science learning in the first years of schooling. An intervention study with sequenced learning material on the topic of ‘floating and sinking'. *International Journal of Science Education*, *36*, 1751–1771.

Lipowsky, F., Rakoczy, K., Pauli, C., Drollinger-Vetter, B., Klieme, E., & Reusser, K. (2009). Quality of geometry instruction and its short-term impact on students' understanding of the Pythagorean theorem. *Learning and Instruction, 19*, 527–537.

Lüdtke, O., Marsh, H. W, Robitzsch, A., & Trautwein, U. (2011). A 2 × 2 taxonomy of multilevel latent contextual models: Accuracy-bias trade-offs in full and partial error-correction models. *Psychological Methods, 16*, 444–467.

Lüdtke, O., Robitzsch, A., Trautwein, U., & Kunter, M. (2009). Assessing the impact of learning environments: How to use student ratings of classroom or school characteristics in multilevel modeling. *Contemporary Educational Psychology, 34*, 120–131.

MacKinnon, D. P. (2008). *Introduction to statistical mediation analysis*. New York: Lawrence Erlbaum Associates.

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

Mahler, D., Großschedl, J., & Harms, U. (2017). Using doubly latent multilevel analysis to elucidate relationships between science teachers' professional knowledge and students' performance. *International Journal of Science Education, 39*, 213–237.

Mansour, N. (2009). Science teachers' beliefs and practices: Issues, implications and research agenda. *International Journal of Environmental & Science Education, 4*, 25–48.

Martin, M. O., Mullis, I. V. S., & Foy, P. (2008). *TIMSS 2007 international science report*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center.

Martin, M. O., Mullis, I. V. S., Foy, P., & Hooper, M. (2016). *TIMSS 2015 international results in science.* Boston: TIMSS & PIRLS International Study Center.

Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning? The case for guided methods of instruction. *American Psychologist, 59*, 14–19.

McCaffrey, D. F., Lockwood, J. R., Koretz, D., Louis, T. A., & Hamilton, L. (2004). Models for value-added modeling of teacher effects. *Journal of Educational and Behavioral Statistics, 29*, 67–101.

Meschede, N., Fiebranz, A., Möller, K., & Steffensky, M. (2017). Teachers' professional vision, pedagogical content knowledge and beliefs: On its relation and differences between pre-service and in-service teachers. *Teaching and Teacher Education*, *66*, 158–170.

Möller, K. & Jonen, A. (2005). Die KiNT-Boxen – Kinder lernen Naturwissenschaft und Technik. Klassenkisten für den Sachunterricht. Paket 1: Schwimmen und Sinken [Materials for teaching the topic of floating and sinking]. Essen, Germany: Spectra.

Muthén, L. K. & Muthén, B. O. (1998–2012). *Mplus user's guide. Seventh edition*. Los Angeles, CA: Muthén & Muthén.

Naumann, A., Hochweber, J., & Hartig, J. (2014). Modeling instructional sensitivity using a longitudinal multilevel differential item functioning approach. *Journal of Educational Measurement, 51*, 381–399.

Odden, T. O. B., & Russ, R. S. (2019). Defining sensemaking: Bringing clarity to a fragmented theoretical construct. *Science Education, 103*, 187–205.

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*, 261–284.

Pianta, R. C. & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: standardized observation can leverage capacity. *Educational Researcher, 38*, 109–119.

Polikoff M. S (2015). The stability of observational and student survey measures of teaching effectiveness. *American Journal of Education, 121*, 183–212.

Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods, 15*, 209–233.

Raudenbush, S. W., & Bryk, A. S. (2010). *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks, CA: Sage.

Rice, D. C. (2005). I didn't know oxygen could boil! What preservice and inservice elementary teachers' answers to 'simple' science questions reveals about their subject matter knowledge. *International Journal of Science Education, 27*, 1059–1082.

Richardson, P. W., Karabenick, S. A., & Watt, H. M. G. (2014). *Teacher motivation: Theory and practice*. New York: Routledge.

Rimm-Kaufman, S. & Hamre, B. (2010). The role of psychological and developmental science in efforts to improve teacher quality. *Teachers College Record, 112*, 2988–3023.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, Schools, and Academic Achievement. *Econometrica, 73*, 417–458.

Roth, G., Assor, A., Kanat-Maymon, Y., & Kaplan, H. (2007). Autonomous motivation for teaching: How self-determined teaching may lead to self-determined learning. *Journal of Educational Psychology, 99*, 761–774.

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

Schmitz, G. S. & Schwarzer, R. (2000). Selbstwirksamkeitserwartung von Lehrern: Längsschnittbefunde mit einem neuen Instrument [Perceived self-efficacy of teachers: Longitudinal findings with a new instrument]. *Zeitschrift für Pädagogische Psychologie, 14*, 12–25.

Schneider, M., Vamvakoussi, X., & van Dooren, W. (2012). Conceptual change. In N. M. Seel (Ed.), *Encyclopedia of the sciences of learning* (pp. 735–738). New York: Springer.

Seidel, T. & Shavelson, R. (2007). Teaching effectiveness research in the past decade. *Review of Educational Research, 77*, 454–499.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*, 4–21.

Shulman, L. S. (1998). Theory, practice, and the education of professionals. *Elementary School Journal, 98*, 511–526.

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

Timperley, H. (2008). *Teacher professional learning and development*. Geneva: International Bureau of Education.

Tsai, C.-C. (2002). Nested epistemologies: Science teachers' beliefs of teaching, learning and science. *International Journal of Science Education, 24*, 771–783.

Tschannen-Moran, M., & Hoy, A. W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education, 17*, 783–805.

Tschannen-Moran, M., Woolfolk Hoy, A., & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research, 68*, 202–248.

van de Pol, J., Volman, M., & Beishuizen, J. (2010). Scaffolding in teacher-student interaction: a decade of research. *Educational Psychology Review, 22*, 271–297.

Vosniadou, S. (2013). Conceptual change in learning and instruction: The framework theory approach. In S. Vosniadou (Ed.), *International Handbook of Research on Conceptual Change* (pp. 11–30). New York: Routledge.

Voss, T., Kleickmann, T., Kunter, M., & Hachfeld, A. (2013). Mathematics teachers' beliefs. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Eds.), *Cognitive activation in the mathematics classroom and professional competence of teachers – results from the COACTIV project* (pp. 249–272). New York: Springer.

Wandersee, J., Mintzes, J., & Novak, J. (1994). Research on alternative conceptions in science. In D. Gabel (Ed.), *Handbook of research on science teaching and learning* (pp. 177–210). New York: Macmillan Publishing Company.

Wang, M. C., Haertel, G. D., & Walberg, H. J. (1993). Toward a knowledge base for school learning. *Review of Educational Research, 63*, 249–294.

Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika, 54*, 427–450.

Warwas, J., Hertel, S., & Labuhn, A. S. (2011). Bedingungsfaktoren des Einsatzes von adaptiven Unterrichtsformen im Grundschulunterricht [Conditions for the application of adaptive teaching in elementary school]. *Zeitschrift für Pädagogik, 57*, 854–867.

Weiß, R. H. (2006). *CFT 20-R. Grundintelligenztest Skala 2 – Revision* [Culture fair test]. Göttingen, Germany: Hogrefe.

Yuan, K. H., & Bentler, P. M. (2000). Three likelihood based methods for mean and covariance structure analysis with nonnormal missing data. *Sociological Methodology, 30*, 165–200.

Zee, M. & Koomen, H. M. Y. (2016). Teacher self-efficacy and its effects on classroom processes, student academic adjustment, and teacher well-being: A synthesis of 40 years of research. *Review of Educational Research, 86*, 981–1015.

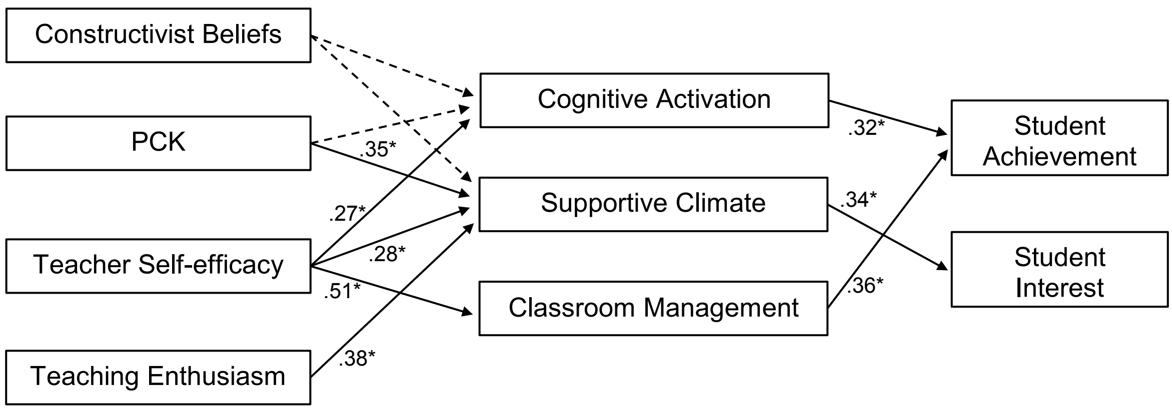

Figure 1: Significant paths of the full mediation model. Dashed lines indicate hypothesized but empirically insignificant relations. Classroom level effects. $*p < .05$, one-tailed test

Table 1

Descriptive Statistics

	N	*M*	*SD*
Teacher/classroom-level variables			
PCK[1]	52	8.75	3.57
Self-efficacy	52	72.10	9.72
Teaching enthusiasm	51	3.48	0.38
Constructivist beliefs	52	3.51	0.39
Cognitive activation (observer)	53	3.20	0.82
Supportive climate (students)[2]	54	3.30	0.31
Classroom management (students)[2]	54	2.56	0.47
Individual-level variables			
Pre-interest	992	3.11	0.90
Post-interest	966	2.85	0.99
Pretest conceptual understanding	987	0.02	0.90
Posttest conceptual understanding	976	0.01	0.95
Science competence	997	-0.32	1.07
Cognitive abilities	991	104.03	14.77

Note. [1]min = 0, max = 16; [2]Classroom aggregates.

Table 2

Correlations at the Classroom Level

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1) PCK												
(2) Self-efficacy	-.25											
(3) Teaching enthusiasm	-.11	.30*										
(4) Constructivist beliefs	-.22	.21	.20									
(5) Cognitive activation	-.11	.29*	-.05	.11								
(6) Supportive climate[2]	.27	.30*	.43*	.18	-.04							
(7) Classroom management[2]	-.10	.42*	.07	.13	.14	.37*						
(8) Post-achievement[2]	-.20	.36*	.01	.08	.38*	.04	.43*					
(9) Post-interest[2]	.24	.25	.30*	.23	-.01	.57*	.36*	.20				
(10) Pre-achievement[2]	.03	-.16	-.26	.21	.02	-.14	-.03	.11	-.19			
(11) Pre-interest[2]	-.17	.32*	.00	.35	.20	.23	.33*	.11	.30*	-.04		
(12) Cognitive abilities[2]	-.00	.29*	-.24	.12	.15	.05	.42*	.45*	.12	.33*	.07	
(13) Science competence[2]	-.21	.19	-.10	-.03	.17	-.17	-.08	.32*	-.23	.24	.04	.54*

Note. *$p < .05$. [2]Classroom aggregates.

Table 3

Multilevel regression analyses: Teacher competence as a predictor of students' achievement and interest after the teaching units

	Main effect of each single predictor		Unique contribution of each predictor	
	Models 1a-d	Models 2a-d	Model 3	Model 4
Dependent Variable	Achievement	Interest	Achievement	Interest
Teacher Competence				
Pedagogical Content Knowl. (PCK)	-.18 (.13)	.25 (.11)*	-.10 (.14)	.36 (.10)*
Constructivist beliefs (CB)	.05 (.14)	.14 (.15)	-.10 (.14)	.08 (.13)
Self-efficacy (SE)	.33 (.15)*	.33 (.14)*	.34 (.18)*	.25 (.13)*
Teaching enthusiasm (TE)	.13 (.15)	.40 (.14)*	-.00 (.17)	.33 (.15)*
R^2 (between) of PCK/CB/SE/TE	.28/.26/.31/.27	.22/.15/.16/.23	.32	.43

Note. Standardized regression weights; standard errors are in parentheses. **p* < .05, †*p* < .10, one-tailed test. PCK = Pedagogical content knowledge, CB = Constructivist beliefs, SE = Self-efficacy, TE = Teaching enthusiasm; Model 1a: PCK, Model 1b: CB, Model 1c: SE, Model 1d: TE, Model 2a: PCK, Model 2b: CB, Model 2c: SE, Model 2d: TE.

We included the following covariates at the individual and at the classroom level of analysis in all models: prior science competence, cognitive abilities, and pretest scores of conceptual understanding or interest, respectively. We also controlled for dummy-coded treatment conditions at the classroom level.

Table 4

Multilevel regression analyses predicting student achievement and interest: multilevel path model

Classroom level predictors	Mediator variables			Dependent variables	
	Cognitive activation	Supportive climate	Classroom management	Interest	Achievement
Teacher Competence					
PCK	-.05 (.15)	.35 (.11)*	.02 (.14)	.22 (.12)	-.11 (.14)
CB	.10 (.13)	.08 (.13)	.00 (.12)	.08 (.13)	-.12 (.13)
SE	.27 (.16)*	.28 (.13)*	.51 (.13)*	.17 (.13)	.17 (.15)
TE	-.14 (.15)	.38 (.11)*	-.07 (.13)	.17 (.15)	.10 (.18)
Teaching Quality					
Cognitive activation	-	-	-	-.06 (.09)	.32 (.13)*
Supportive climate	-	-	-	.34 (.17)*	-.09 (.17)
Classroom management	-	-	-	.00 (.16)	.36 (.14)*
R^2	.14	.47	.31	.68	.50

Note. Standardized regression weights; standard errors are in parentheses. *$p < .05$, one-tailed test. PCK = Pedagogical content knowledge, CB = Constructivist beliefs, SE = Self-efficacy, TE = Teaching enthusiasm.

We included the following covariates at the individual and at the classroom level of analysis in this model: prior science competence, cognitive abilities, and pretest scores of conceptual understanding or interest, respectively. We also controlled for dummy-coded treatment conditions at the classroom level.