Longitudinal relations between teaching-related motivations and student-reported teaching quality

Anna-Katharina Praetorius [a, *], Fani Lauermann [b], Robert M. Klassen [c], Oliver Dickhäuser [d], Stefan Janke [d], Markus Dresel [e]

[a] German Institute for International Educational Research (DIPF), Germany
[b] Department of Psychology, University of Bonn, Germany
[c] Department of Education, University of York, United Kingdom
[d] Department of Educational Psychology, University of Mannheim, Germany
[e] Department of Psychology, University of Augsburg, Germany

1. Introduction

Teaching-related motivations constitute a core element of teachers' professional competence, and are assumed to influence such important outcomes as teachers' instructional practices and teaching quality (e.g., Kunter et al., 2013; Zee & Koomen, 2016). Available research generally supports positive associations between aspects of teacher motivation and teaching characteristics such as autonomy support or monitoring (e.g., Hein et al., 2012; Kunter et al., 2008; Morris-Rothschild & Brassard, 2006; Pelletier, Séguin-Lévesque, & Legault, 2002; Roth, Assor, Kanat-Maymon, & Kaplan, 2007). However, the vast majority of this research is cross-sectional and thus potential longitudinal reciprocal links between teaching-related motivations and teaching quality remain relatively unexplored (see also Soodak & Podell, 1998). This constitutes an important gap in the literature, because argumentation for the high relevance of teacher motivation regularly refers to its longitudinal effects on teaching quality, and cross-sectional relations are not sufficient to support the existence of such effects. Instead, there could be no longitudinal relation between these aspects at all (e.g., because both depend on a third variable), there might be reciprocal links, or longitudinal influences might in fact be in the opposite direction than previously assumed (Kunter & Holzberger, 2014).

* Corresponding author. German Institute for International Educational Research (DIPF), Schlossstrasse 29, 60686 Frankfurt, Germany.
E-mail address: praetorius@dipf.de (A.-K. Praetorius).

Indeed, recent evidence suggests that teacher motivation is not only a predictor of teaching quality (as is typically assumed in the extant literature), but is also influenced by teachers' prior classroom experiences and quality of teaching. Specifically, Holzberger, Philipp, and Kunter (2013) demonstrated that two dimensions of student-perceived teaching quality (cognitive activation and learning support) had a positive longitudinal predictive effect on teachers' self-efficacy whereas no significant predictive effects of teachers' self-efficacy on student-perceived teaching quality were found.

Disentangling potential reciprocal links between teacher motivation and teaching quality is important for several reasons. For instance, gaining a more advanced understanding of the longitudinal relations between aspects of teacher motivation and teaching quality has implications for teacher training and professional development; if teacher motivation has a considerable effect on teaching quality, it might be useful to not only aim at enhancing teaching quality directly, but also indirectly through changing teachers' motivations (for a similar argument regarding students, see Wigfield & Eccles, 2000). Analogously, if teacher motivation is primarily a consequence of their classroom experiences (e.g., mastery experiences with high quality teaching), then this might be a key pathway towards improving teachers' professional wellbeing. Finally, if these two types of constructs do not significantly predict each other over time, but are nevertheless correlated within each time point, research attention should be devoted to third variables that might shape both teachers' motivations and instructional quality (e.g., professional knowledge, prior training, and teaching beliefs). In the following sections, we discuss the role of teaching-related motivations in the instructional process, conceptualizations of teaching quality, and possible longitudinal relations between teachers' motivations and teaching quality.

1.1. *Aspects of teacher motivation: definition and relevance*

The term motivation generally refers to the underlying reasons behind people's actions (Graham & Weiner, 1996). Because these reasons can be very diverse, motivation is an umbrella term for a variety of internal characteristics and processes. Several theories have been developed that differentiate types of motivations. One of the most prominent frameworks is expectancy-value theory (Eccles, 2009). It proposes that achievement-related behaviors can be predicted by individuals' beliefs about whether they can carry out relevant actions successfully (expectancy component) as well as the value they attach to these actions and expected results (value component). Teachers' self-efficacy (i.e., the belief in one's own capabilities) and teachers' enthusiasm for teaching (i.e., intrinsic value seen in teaching) can be seen as pivotal representations of these two basic motivational constructs; self-efficacy is closely related to the expectancy component of motivation, enthusiasm to the value component. Relating self-efficacy and enthusiasm to the logic of the expectancy-value framework of motivation indicates that core aspects of teacher motivation can be captured by investigating self-efficacy and enthusiasm, because each of them represents a central aspect of human motivation. Due to their critical role for teachers and teaching (Kunter, 2011), these two constructs have attracted substantial attention in research on teacher motivation. For instance, both self-efficacy and enthusiasm for teaching have been linked to such important teacher outcomes as burnout (e.g., Kunter, Frenzel, Nagy, Baumert, & Pekrun, 2011; Skaalvik & Skaalvik, 2007) and job satisfaction (e.g., Caprara, Barbaranelli, Borgogni, & Steca, 2003; Kunter et al., 2011; Vieluf, Kunter, & van de Vijver, 2013). Teachers' self-efficacy in particular has been identified as by far the most frequently studied aspect of teacher motivation (Woolfolk Hoy, 2008).

Teachers' self-efficacy reflects a belief in teachers' own capabilities to influence student learning and to manage the learning environment (Klassen & Chiu, 2010; Dicke, Parker, Holzberger, Kunter, & Leutner, 2015; Tschannen-Moran & Woolfolk Hoy, 2001). Self-efficacy constitutes a motivational construct, because individuals would be unlikely to engage in activities or to pursue goals that they believe might exceed their capabilities; conversely, efficacious individuals are more likely than less efficacious ones to set challenging goals, to persist in the face of difficulty, and to show resilience in the face of failure (Bandura, 1997). Drawing on Bandura's (1997) socio-cognitive theory, Tschannen-Moran, Woolfolk Hoy, and Hoy (1998) proposed that teachers' self-efficacy develops cyclically. Efficacy-building experiences (e.g., mastery experiences such as producing or failing to produce desired classroom outcomes) affect teachers' perceived teaching competence and thus their sense of self-efficacy. Teachers' self-efficacy, in turn, influences subsequent levels of performance, mediated via teachers' goals, effort, and persistence. Teachers' performance provides efficacy-relevant information, therefore starting a new cycle of self-efficacy-building experiences and judgments.

Teachers' enthusiasm refers to an affective, inner-personal state that can be categorized as both a positive emotion and an intrinsic type of motivation (Kunter et al., 2011).[1] Accordingly, teacher enthusiasm is investigated in research on both teachers' emotions (see e.g., Frenzel, Goetz, Lüdtke, Pekrun, & Sutton, 2009, labeled as teacher enjoyment) and motivation (see e.g., Kunter et al., 2011). Two components of teacher enthusiasm have emerged in motivation research: enthusiasm for the subject matter taught by the teacher, and enthusiasm for teaching. Only the latter has been found to be positively linked to students' perceptions of teaching quality (Kunter et al., 2008). In a comprehensive review of the literature, Kunter and Holzberger (2014) proposed that teacher enthusiasm represents an intrinsic orientation towards teaching that is influenced by school characteristics (e.g., school climate), teacher characteristics (e.g., self-efficacy), and student characteristics (e.g., achievement), and influences teacher characteristics (e.g., well-being), teaching quality (e.g., autonomy support for students), and student outcomes (e.g., achievement).

A common assumption in research on both teachers' self-efficacy and enthusiasm for teaching is that such motivational factors matter due to their effects on teachers' behaviors, which, in turn, can influence students' motivation and achievement (e.g., Klassen & Tze, 2014; Kunter et al., 2013; Tschannen-Moran & Woolfolk Hoy, 2001; Tschannen-Moran et al., 1998; Ware & Kitsantas, 2007). One of the most important proximal outcomes of teacher motivation within this functional chain is teachers' instructional practices and their teaching quality. The main objective of the present study was therefore to examine the longitudinal relations between crucial teaching-related motivations (teachers' self-efficacy and enthusiasm for teaching) and dimensions of teaching quality.

1.2. *Teaching quality: conceptualization and measurement*

Teaching quality is one of the key factors influencing student learning over and above the effects of student characteristics (see review in Hattie, 2009). In the context of teacher effectiveness research (see review in Seidel & Shavelson, 2007) teaching quality

[1] Enthusiasm is in some contexts also conceptualized as a teaching style (see e.g., Patrick, Turner, Meyer, & Midgley, 2003). In the present study, we use Kunter et al.'s (2011) conceptualization, according to which enthusiasm reflects a subjective experience and has motivational implications.

is defined as teaching characteristics that lead to an enhancement of student characteristics, mainly focusing on student achievement. Different sets of such characteristics have been proposed and several attempts have been made to integrate separate notions of teaching quality into an overarching model. Interestingly, researchers from different cultural and educational contexts, such as Germany and the United States, have identified similar instructional quality dimensions (see Fauth, Decristan, Rieser, Klieme, & Büttner, 2014; Kunter et al., 2013; Lipowsky et al., 2009; Pianta & Hamre, 2009; Reyes, Brackett, Rivers, White, & Salovey, 2012). Three generic teaching quality dimensions have been proposed that are assumed to be essential for high quality teaching in different education systems, school types, grade levels, and school subjects (see Klieme & Rakoczy, 2003): classroom management (also labeled classroom organization), learning support (also labeled emotional support), and cognitive activation (also labeled instructional support).

Classroom management is a characteristic of teaching quality that has gained much attention for several decades since the initial work by Kounin (1970). Classroom management focuses on maximizing students' learning time (i.e., time on task) by preventing or by dealing effectively with disruptions and disciplinary conflicts. Ways to achieve high quality classroom management include, for instance, clearly explicated and consistently implemented rules and routines and efficient classroom organization. Classroom management has been linked to enhanced student achievement as well as student motivation (Fauth et al., 2014; Kunter, 2005; Lipowsky et al., 2009; Rakoczy, 2008).

Learning support refers to teachers' attempts to account for the needs and the perspectives of their students in the instructional process (Cornelius-White, 2007; Davis, 2003; Pianta & Hamre, 2009). This dimension is closely aligned with and derived from self-determination theory (Ryan & Deci, 2000) and focuses on fostering students' experiences of competence, autonomy, and social relatedness. The dimension reflects, for instance, a constructive way of dealing with student errors, constructive feedback, student-oriented individual support, and positive teacher-student relationships. Learning support has been shown in the literature to be positively linked to enhanced student motivation (e.g., Fauth et al., 2014; Kunter et al., 2013; Lipowsky et al., 2009; Rakoczy, 2008).

Finally, cognitive activation aims at assisting students' higher-level thinking (see, e.g., the concept of teaching for understanding, Cohen, 1993; Mayer, 2004). It is based on constructivist learning theories (e.g., Dewey, 1916); cognitively activating instruction utilizes challenging tasks and questions that elicit students' deep-level thinking, activates prior knowledge and initiates content-related discourse. Cognitive activation has been linked to higher student achievement (Baumert et al., 2010; Fauth et al., 2014; Lipowsky et al., 2009).

There are different approaches to the assessment of teaching quality, including observations by independent evaluators, teacher self-reports, and student reports. Each of these approaches has its advantages and disadvantages (see e.g., Clausen, 2002; Kunter & Baumert, 2006). In the present study, we rely on student ratings of their teachers. We chose this approach, because it allows us to avoid a so called common method bias (see, e.g., Williams, Hartman, & Cavazotte, 2010); assessing the associations between different constructs (here, teacher motivation and teaching quality) from the perspective of the same source (teachers) can lead to inflated estimates. Using students' ratings of teaching quality thus provides a more rigorous test of associations. Compared to observer ratings, student ratings allow a more general, long-term view on teaching, because external observers usually observe only one or a few lessons, which is problematic if lesson quality varies substantially (e.g., Praetorius, Pauli, Reusser, Rakoczy, & Klieme, 2014). In addition, relative to other indicators of teaching quality, students' perceptions are more proximal to student-related outcomes such as student achievement (e.g., Clausen, 2002).

1.3. Relations between teacher motivation and teaching quality

Studies of the cross-sectional associations between teacher self-efficacy and dimensions of teaching quality have produced mixed results (see meta-analysis by Zee & Koomen, 2016): Studies that show significant positive relations between teacher self-efficacy and teaching quality (e.g., Holzberger et al., 2013) seem to be just as common as studies showing no associations at all (e.g. Jamil, Downer, & Pianta, 2012). Regarding enthusiasm for teaching, Kunter et al. (2008) as well as Holzberger, Phillip, and Kunter (2016) investigated its relation with three dimensions of teaching quality (classroom management, learning support, and cognitive activation) in cross-sectional studies. All three dimensions (measured via student ratings or student teachers' self-reports) were significantly related to enthusiasm for teaching.

Because the vast majority of existing studies are cross-sectional, the directionality of the investigated associations is uncertain, with three main possibilities (see Fig. 1). First, as stated previously, teacher motivation is typically conceptualized as an antecedent of teachers' behaviors and approaches to teaching. The general mechanisms are assumed to be that more relative to less motivated teachers (a) behave differently in the classroom, for instance, by investing more effort in teaching, working harder, setting more ambitious goals, and showing higher persistence as well as enhanced concentration and attention in their instruction (see Holzberger, Philipp, & Kunter, 2014; Kunter & Holzberger, 2014; Tschannen-Moran et al.,1998); and (b) are more willing to engage and invest effort towards professional development activities (see Klassen, Tze, Betts, & Gordon, 2011; Lohman, 2006; Ross & Bruce, 2007). Accordingly, teacher motivation could lead to higher teaching quality.

Second, individuals' motivation is shaped by prior experiences of success or failure in achievement situations (e.g., Eccles & Wigfield, 2002). For teachers, experiences of success or failure might refer to their perceived level of instructional quality. The higher teachers perceive their teaching quality, the more confident they should be regarding their teaching abilities (self-efficacy) and the more enthusiastic regarding teaching (enthusiasm for teaching). Thus, higher levels of perceived teaching quality might foster higher levels of teacher motivation.

Third, further characteristics (e.g., class characteristics such as the mean achievement level of the students) and teacher characteristics (e.g., professional knowledge) might not only shape their quality of teaching but also teachers' motivations. Thus, the relations between teaching quality and teacher motivation could dependent on third variables, in addition to their potentially reciprocal links.

Empirical investigations regarding which of these types of relations is dominant for the relation between teacher motivation and teaching quality are scarce, because very few longitudinal studies of these associations exist. Specifically, using longitudinal analyses, Holzberger et al. (2013) found no effect of self-efficacy on student-perceived teaching quality. Instead, two dimensions of teaching quality (cognitive activation and learning support) had a positive longitudinal predictive effect on teachers' self-efficacy. Longitudinal analyses of the potential reciprocal links between enthusiasm and teaching quality have not been conducted (see review in Kunter & Holzberger, 2014). However, analogous to teachers' self-efficacy, reciprocal effects between enthusiasm for teaching and teaching quality are plausible. Potential dependencies on third variables have not been discussed or empirically

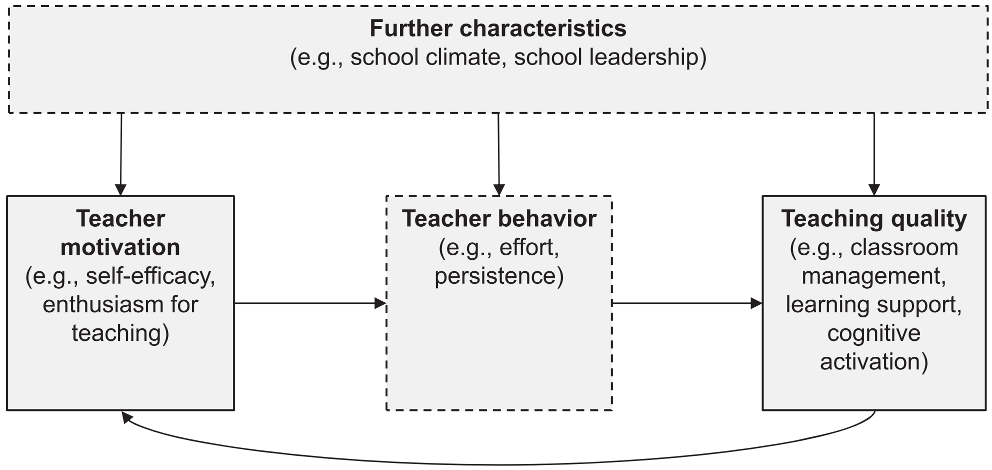

Fig. 1. Assumptions regarding the influences and mediating processes for teacher motivation and teaching quality. Solid lined boxes indicate aspects that have been measured in the present study; dashed lined boxes indicate aspects that are hypothesized based on the literature but are not investigated in the present study.

investigated so far, neither for teachers' self-efficacy, nor for their enthusiasm for teaching.

1.4. Conditions for identifying longitudinal relations between teacher motivation and teaching quality

Whether or not studies succeed at identifying longitudinal effects in cross-lagged analyses, depends at least on two factors, (a) the chosen time lag for the analyses (see e.g., Dormann & Griffin, 2015; Voelkle, Oud, Davidov, & Schmidt, 2012), and (b) the consideration of trait-like individual differences (Hamaker, Kuiper, & Grasman, 2015).

The length of the time lag for analyses of reciprocal links between two given constructs must be carefully chosen to match the expected time frame of influence between these constructs. However, the decision which time lag to choose is often not based on sound theoretical or methodological evidence. According to Dormann and Griffin (2015), time lags used in psychological research are often too long, so that potentially existing cross-lagged effects are unlikely to be detected. Consistent with this assumption, Holzberger et al. (2013) proposed that the lack of significant cross-lagged effects of teachers' self-efficacy on student-reported teaching quality in their study could be at least partially attributable to the one-year measurement interval used in their study, which might have been too long to detect such effects. Shorter-term effects are plausible due to the cyclical nature of self-efficacy (Bandura, 1997), which is continuously influenced by teachers' subjective experiences of success or failure in the classroom. Due to changes in curriculum, learning goals, or possible developmental changes in their students, teachers' classroom experiences and teaching quality during a given school year should be more relevant for their self-efficacy ratings that school year than for their ratings in following years (for a similar argument regarding the relevance of short-term effects in psychological research, see Dormann & Griffin, 2015). Therefore, not only long-term reciprocal effects (e.g., across school years), but also shorter-term reciprocal effects (e.g., within the same school year) between teaching quality and teacher self-efficacy should be considered. An analogous rationale applies regarding the appropriate time lags for cross-lagged analyses of enthusiasm for teaching and teaching quality, although no prior cross-lagged analyses exist that could serve as a reference point.

Hamaker et al. (2015) discussed an additional challenge associated with the traditional cross-lagged panel approach, namely trait-like individual differences. For example, analyses of teaching efficacy and teaching quality over time are likely influenced by trait-like differences between teachers (i.e., some teachers are consistently more efficacious than others and might consistently provide higher quality instruction than others). Analyses of reciprocal influences examine whether changes in self-efficacy over time correspond to changes in teaching quality, but generally fail to account for stable trait-like associations between these constructs. Specifically, the autoregressive paths that aim to account for temporal stability in traditional cross-lagged panel models, implicitly assume that all people vary over time around the same means of the characteristics under investigation. Because this assumption is unlikely to hold true (individuals can have different means), the existing, but not taken into account trait-like differences can lead to biased estimates of the cross-lagged paths. Taking such trait-like differences into account allows disentangling stable relations between constructs and actual influences over time. The importance of trait-like associations over time is also evident in the relatively large test-retest correlations for teachers' self-efficacy found in Holzberger et al.'s (2013) study, based on a one-year period ($r_{tt} = 0.84$). No empirical evidence regarding the stability of teachers' enthusiasm for teaching exists so far.

1.5. Research questions and hypotheses

The present study was designed to examine the reciprocal links between teaching-related motivations (self-efficacy and enthusiasm for teaching) and teaching quality (student-reported classroom management, learning support, and cognitive activation). Based on theoretical assumptions in the literature on self-efficacy, positive effects of self-efficacy on teaching quality were expected (Hypothesis 1.1). Consistent with Holzberger et al. (2013), we expected significant positive effects of student-reported teaching quality on teachers' subsequent self-efficacy (Hypothesis 1.2). We further expand upon earlier evidence by examining the cross-lagged associations between teachers' enthusiasm and student-reported teaching quality. Analogous hypotheses to the ones for self-efficacy were examined for teachers' enthusiasm for teaching, expecting positive links from enthusiasm for teaching to teaching quality (Hypothesis 2.1), as well as positive cross-lagged paths from student-reported teaching quality to subsequent teacher enthusiasm (Hypothesis 2.2).

We additionally extend earlier research through the following research objectives: First, we investigate cross-lagged effects between teacher motivation and teaching quality across two different time lags (one year and six months), so that it is possible to examine not only long-term, but also shorter-term effects. Based on the rationale presented by Holzberger et al. (2013), we expected stronger cross-lagged effects for the shorter time lag than for the longer time lag (Hypothesis 3). Second, in contrast to earlier research, our analyses account for stable inter-individual differences in teacher motivation and teaching quality (see Hamaker et al., 2015), in order to separate potential cross-lagged effects from trait-like associations between these two constructs. We expected that motivational orientations are relatively stable traits rather than situation-specific measures (Hypothesis 4); thus, taking into consideration trait-like associations should lead to a decrease in the likelihood of finding cross-lagged effects. Third, we conduct optimal time lag analyses to estimate the most appropriate time frame for future longitudinal analyses of the associations between teacher motivation and teaching quality (Dormann & Griffin, 2015).

2. Method

2.1. Sample and procedure

A total of 288 academic-track secondary schools ("Gymnasien") in the German federal state of Baden-Württemberg were invited to participate in this study, 57 of which agreed to participate. Thirteen schools were located in urban areas and 44 in rural areas; 46 of the schools were public and 11 private. The number of teachers per school ranged between one and five, because only mathematics teachers teaching in 5th grade classrooms in the school year 2011/12 were invited to participate. We restricted the study to 5th grade classrooms to ensure comparability across classrooms. Data from all participating teachers were included in the analyses across three measurement points. A total of 165 German mathematics teachers (57% female, mean age 41.14 years, $SD = 13.44$, average teaching experience 13.30, $SD = 12.29$ with a range between 0 and 40 years) and their 4273 students (50% female, mean age 13.88 years, $SD = 0.47$) participated at Time 1.

The data collection was continued in the 2012/13 school year for those teachers who taught the same class the following year. This led to a reduction of the targeted sample size to 70 teachers and 1538 students at Time 2, and 69 teachers and 1483 students at Time 3. Of these, the data of 68 (i.e., a response rate of 97%) and 42 (i.e., a response rate of 69%) teachers were available for the analyses at Times 2 and 3, respectively. All available data were included in the analyses in order to utilize the maximum available information. The full information maximum likelihood (FIML) estimator was used to handle missing data (e.g., Arbuckle, 1996). FIML is adequate for multilevel data even with a large amount of missing data (Enders, Mistler, & Keller, 2016), if missing data is at least missing at random. As shown in Table 1, a dummy variable indicating whether a teacher had complete vs. incomplete data was unrelated to any other variables of interest, which suggests that teachers who participated at all three time points did not differ significantly from teachers with incomplete data.

Teachers and students were surveyed three times with a one year and a six months interval (November 2011, November 2012, and June 2013). The first time interval is comparable to Holzberger et al. (2013) with respect to its length (Time 1 to Time 2); the second time interval was included to examine potential shorter-term effects within the same school year (Time 2 to Time 3; cf. Dormann & Griffin, 2015).

2.2. Instruments

Teacher measures. For the sake of comparability, teacher measures used in the present study were informed by prior evidence on the associations between teacher motivation and student-reported teaching quality (see Holzberger et al., 2013, for teacher self-efficacy; and Kunter et al., 2008 for enthusiasm for teaching).

Teachers' self-efficacy. Teacher self-efficacy was assessed with a 10-item scale developed by Schwarzer and Schmitz (1999). The scale is widely used in German-speaking countries and has been validated with diverse national and international teacher samples (e.g., Schmitz & Schwarzer, 2000). The scale covers a broad range of aspects relevant for the teaching profession (e.g., working with students, parents, and colleagues). A sample item (translated from German) is: "I am confident that I can develop creative ideas for changing unfavorable instructional structures," rated on a scale from 1 (*disagree*) to 4 (*agree*). Cronbach's α ranged between 0.73 and 0.75 across the three measurement points.

Teachers' enthusiasm for teaching. A scale from the study "Professional competence of teachers, cognitively activating instruction, and the development of students' mathematical literacy" (COACTIV; Kunter et al., 2011) was used to measure teachers' enthusiasm for teaching. The scale includes two items and has shown good predictive validity in relation to both teaching quality indicators and student outcomes (Kunter et al., 2008, 2013). Cronbach's α ranged between 0.66 and 0.74 across the three measurement points. A sample item is: "I teach mathematics in this class with great enthusiasm", rated on a scale from 1 (*strongly disagree*) to 5 (*strongly agree*).

Student measures of teaching quality. The students' ratings of teaching quality were assessed with measures that have been validated across diverse student samples and have shown very good psychometric properties on the class level (see e.g., Kunter & Baumert, 2006; Wild, 1999).

Classroom management. The quality of classroom management was assessed with a three-item scale from the Programme for International Student Assessment (PISA) 2003 survey (Ramm et al., 2006). Cronbach's α ranged between 0.91 and 0.94 across the three measurement points. A sample item is: "In mathematics, it takes a very long time at the start of the lesson until the students settle down and start working," rated on a scale from 1 (*strongly disagree*) to 6 (*strongly agree*).

Learning support. Learning support was assessed with five items from a scale developed by Wild (1999). Cronbach's α ranged between 0.92 and 0.95. A sample item is: "In mathematics, I feel accepted and supported by my teacher," on a scale from 1 (*strongly disagree*) to 6 (*strongly agree*).

Cognitive activation. Cognitive activation was assessed with six items from the PISA 2003 survey (Ramm et al., 2006). Cronbach's α varied between 0.82 and 0.92. A sample item is: "In mathematics, our teacher asks questions that cannot be answered directly but stimulate thinking about them," rated on a scale from 1 (*never*) to 5 (*always*).

2.3. Analyses

The items of each scale were averaged to derive one manifest variable for each construct of interest; these manifest variables were used in subsequent analyses.[2] To account for the nested structure of the data (students within classrooms), two-level models were estimated. The individual student ratings of teaching quality were included on level one (within-class level); the class-aggregated student ratings of teaching quality as well as the

[2] The use of latent variables with multiple item indicators posed problems with model convergence due to an insufficient sample of teachers.

Table 1
Descriptive statistics as well as bivariate correlations between the teacher motivation and the teaching quality scales.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Time 1															
1. Teacher self-efficacy	3.04 (0.33)														
2. Teacher enthusiasm	0.28**	4.21 (0.38)													
3. Classroom management (ICC1 = 0.22, ICC2 = 0.88)	0.00	0.36**	3.13 (0.58)												
4. Learning support (ICC1 = 0.11, ICC2 = 0.75)	0.23**	0.23**	0.22**	4.11 (0.40)											
5. Cognitive activation (ICC1 = 0.03, ICC2 = 0.44)	0.09	0.15	0.15	0.40**	3.45 (0.17)										
Time 2															
6. Teacher self-efficacy	0.64***	0.21	0.06	0.12	0.26	2.99 (0.33)									
7. Teacher enthusiasm	0.37**	0.48**	0.37**	0.43**	0.06	0.25*	4.13 (0.42)								
8. Classroom management (ICC1 = 0.28, ICC2 = 0.91)	0.13	0.20*	0.67**	0.13	0.21	0.07	0.32*	3.34 (0.71)							
9. Learning support (ICC1 = 0.18, ICC2 = 0.85)	0.10	0.22	0.34**	0.43**	0.17	0.30*	0.27*	0.39**	3.81 (0.57)						
10. Cognitive activation (ICC1 = 0.11, ICC2 = 0.75)	−0.02	0.03	0.33**	0.19	0.27*	0.14	0.28*	0.43**	0.69**	3.36 (0.28)					
Time 3															
11. Teacher self-efficacy	0.80***	0.39*	−0.01	0.29	0.08	0.79***	0.36*	0.03	0.18	−0.08	3.01 (0.30)				
12. Teacher enthusiasm	0.17	0.45**	0.25	0.32*	−0.10	0.27	0.67**	0.26	0.40**	0.03	0.18	4.04 (0.40)			
13. Classroom management (ICC1 = 0.32, ICC2 = 0.92)	0.09	0.30*	0.67**	0.13	0.16	0.04	0.29*	0.82**	0.32**	0.36**	0.03	0.30*	3.51 (0.82)		
14. Learning support (ICC1 = 0.19, ICC2 = 0.85)	0.11	0.45**	0.27*	0.57**	0.20	0.18	0.37**	0.34**	0.74**	0.47**	0.25*	0.39*	0.41**	3.61 (0.60)	
15. Cognitive activation (ICC1 = 0.12, ICC2 = 0.77)	0.05	0.31*	0.27*	0.42**	0.27*	0.21	0.43**	0.39**	0.61**	0.66**	0.17	0.31*	0.44**	0.77**	3.40 (0.31)
16. Complete data	−0.07	0.08	−0.08	0.03	0.01	−0.01	−0.06	−0.09	0.07	0.06	0.00	−0.00	−0.11	0.06	0.05

Note. In the diagonal, mean values and, in parentheses, standard deviations of the variables are displayed. Measurement intervals were 12 months (t1−t2) and 6 months (t2−t3) respectively. Complete data is coded as 1 (complete) and 0 (incomplete); n_{t1} = 168 classes with 4273 students and 168 teachers; n_{t2} = 70 classes with 1538 students and 68 teachers; n_{t3} = 69 classes with 1438 students and 42 teachers.
$^{**}p < 0.01$; $^{*}p < 0.05$.

Table 2
Measurement invariance tests over time for the teacher motivation and teaching quality constructs.

Model	χ^2	*df*	*p*	CFI	RMSEA
Self-efficacy					
Configural invariance	40.10	24	0.02	0.92	0.07
Metric invariance	43.00	28	0.03	0.92	0.06
Difference between models	$\Delta\chi^2$ ($\Delta df = 4$) = 2.9, $p = 0.57$				
Enthusiasm					
Configural invariance	3.76	6	0.71	1.00	0.00
Metric invariance	5.23	8	0.73	1.00	0.00
Difference between models	$\Delta\chi^2$ ($\Delta df = 2$) = 1.47, $p = 0.48$				
Classroom management					
Configural invariance	63.97	24	0.00	0.96	0.10
Metric invariance	68.22	28	0.00	0.96	0.09
Difference between models	$\Delta\chi^2$ ($\Delta df = 4$) = 4.25, $p = 0.37$				
Learning support					
Configural invariance	56.60	24	0.00	0.96	0.09
Metric invariance	64.07	28	0.00	0.95	0.10
Difference between models	$\Delta\chi^2$ ($\Delta df = 4$) = 7.47, $p = 0.11$				
Cognitive activation					
Configural invariance	44.76	24	0.01	0.96	0.07
Metric invariance	51.64	28	0.00	0.95	0.07
Difference between models	$\Delta\chi^2$ ($\Delta df = 4$) = 6.88, $p = 0.14$				

Note. CFI = comparative fit index; RMSEA = root-mean square error of approximation. Configural invariance means that the factor structure is constrained to be invariant over time; metric invariance means that factor loadings are constrained to be equal over time.

teacher motivation measures were included on level two (between-class level).

Even though our final models relied on observed variables, the measurement model parts of our final models were tested with latent variables to ensure measurement invariance over time (see McArdle, 2009); thus, for every construct, correlated confirmatory factor analyses including all time points were estimated to test whether a model with metric measurement invariance (i.e., constraining all factor loadings to be equal over time) holds. For constructs with more than 4 items, we used three parcels with randomly assigned items. Model comparisons using chi-square difference tests confirmed that metric invariance holds across all three time points for all constructs (see Table 2).

On level two, cross-lagged panel models were used to examine the reciprocal links between teachers' motivations and class-aggregated student-reported teaching quality across the three time points of the study. These models include autoregressive paths from time point to time point for the teacher motivation variables and teaching quality as well as cross-lagged paths from self-efficacy/enthusiasm at a certain time point to teaching quality at the subsequent time point and vice versa. Specifically, cross-lagged predictive effects of self-efficacy/enthusiasm on teaching quality were tested across all time points (Hypotheses 1.1/Hypothesis 2.1) as well as analogous cross-lagged predictive effects of teaching quality on self-efficacy/enthusiasm (Hypotheses 1.2/Hypothesis 2.2). Additionally, cross-lagged paths from time one to time two were compared to cross-lagged paths from time two to time three to see whether stronger cross-lagged effects occur for the shorter time lag than for the longer time lag (Hypothesis 3).

Two random intercepts were included to account for trait-like inter-individual differences in teaching quality and teacher motivation respectively (Hypothesis 4; see Fig. 2; for further information see Hamaker et al., 2015). Equivalence of the cross-lagged paths was tested by imposing model constraints and comparing the model fit of the constrained and unconstrained models using Satorra-Bentler scaling-corrected chi-square difference tests.

The models were estimated with Mplus 7.11 (Muthén & Muthén, 1998–2012) using the restricted maximum likelihood estimation approach and the FIML estimator. Separate models were examined for each pair of teacher motivation constructs (self-efficacy and

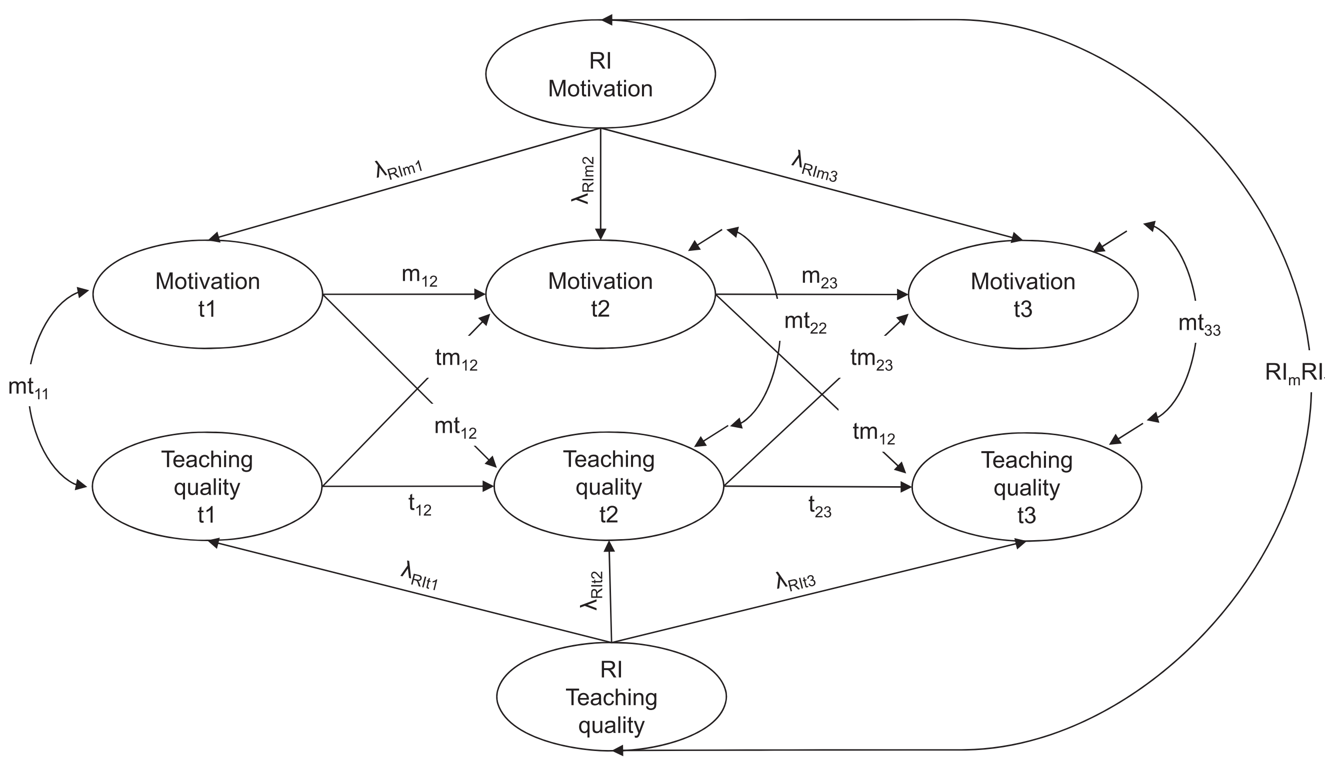

Fig. 2. Random intercept latent cross-lagged panel model for teaching quality and teacher motivation with three measurement points. RI = random intercept; t1 = time point 1; t2 = time point 2; t3 = time point 3; m = motivation; t = teaching quality.

Table 3
Classical cross-lagged panel analyses of the associations between teacher motivation and teaching quality.

Variables		Bivariate correlations r (m, t)		Auto-regressive coefficients		Cross-lagged coefficients	
(1) m	(2) t	(3) Time 1	(4) Time 2	(5) m → m	(6) t → t	(7) m → t	(8) t → m
12 month time interval							
Teacher self-efficacy	Classroom management	0.00	−0.07	0.64***	0.78***	0.06	0.06
	Learning support	0.23**	0.34*	0.65***	0.48***	−0.01	−0.04
	Cognitive activation	0.09	−0.01	0.63***	0.39	−0.04	0.25
Enthusiasm for teaching	Classroom management	0.39***	0.03	0.44***	0.75***	0.08	0.24*
	Learning support	0.27**	−0.12	0.39*	0.46***	0.01	0.28
	Cognitive activation	0.23*	0.55***	0.50***	0.41*	−0.13	0.07
6 month time interval							
Teacher self-efficacy	Classroom management	0.06	0.20	0.79***	0.90***	−0.02	0.08
	Learning support	0.30*	0.15	0.75***	0.80***	−0.03	0.11
	Cognitive activation	0.19	0.16	0.79***	0.78***	0.15	−0.06
Enthusiasm for teaching	Classroom management	0.35**	0.13	0.70***	0.90***	0.00	0.00
	Learning support	0.32*	−0.03	0.60***	0.73***	0.18*	0.32*
	Cognitive activation	0.46***	0.31	0.80***	0.68***	0.26*	−0.21

Note. m = motivation (teacher self-efficacy or enthusiasm for teaching); t = teaching quality indicator (student-reported classroom management, learning support, or cognitive activation). Columns (1) and (2) indicate which variables were included in each cross-lagged panel analysis. Columns (3) and (4) indicate the correlations between these variables at Time 1 (Column 3) and Time 2 (Column 4). Columns (5) and (6) indicate the estimated autoregressive paths for each variable (i.e., its stability) across time points. Column (7) indicates cross-lagged effects of motivation at Time 1 on teaching quality at Time 2, whereas Column (8) shows the analogous cross-lagged effects of teaching quality at Time 1 on motivation at Time 2.
***$p < 0.001$; **$p < 0.01$: *$p < 0.05$. All p levels are reported one-tailed.

enthusiasm) and student-reported teaching quality dimensions (classroom management, learning support, and cognitive activation) as the sample size did not allow including all variables in a single model.

Analyses of the optimal time lag for the presented cross-lagged models were conducted based on suggestions by Dormann and Griffin (2015) for models with reciprocal effects and using the Time 1 to Time 2 lag from the random intercepts latent cross-lagged model, thus controlling for the inter-individual stability of the included variables and for potentially relevant unmeasured third-variables (see Dormann & Griffin, 2015). Their approach seeks to estimate the time lag for which cross-lagged effects are expected to have their maximum values. For doing so, a complex algebraic solution is calculated using the stability coefficients as well as the two cross-lagged path coefficients; the calculation can be done using ordinary least square regression analysis. The formula used can be found in Appendix A.

3. Results

3.1. Descriptive analyses

Descriptive information for all measures as well as bivariate correlations are presented in Table 1. The test-retest correlations ranged between 0.45 and 0.80 for the teacher motivation variables and between 0.27 and 0.82 for the student-reported teaching quality aspects.

The ICC(1) for the student-reported teaching quality ratings ranged between 0.11 and 0.32, with the exception of cognitive activation at the first measurement point, which was 0.03. This means that between 3 and 32 percent of the total variance in these variables is attributable to systematic between-class rather than within-class differences. The ICC(2), a measure of reliability on the class level, was satisfactory for all measures (ranging from 0.75 to 0.92), thus indicating that aggregation on the class level reveals a meaningful class-level construct (see LeBreton & Senter, 2008). The only exception was cognitive activation at the first measurement point (0.44; for a discussion of this finding, see section 4.2).

The cross-sectional correlations between teachers' enthusiasm for teaching and student-rated teaching quality were positive and small to medium sized, ranging from 0.15 to 0.39 (using conventions proposed by Cohen, 1993). These correlations were significant, with the exception of the association between enthusiasm and cognitive activation at Time 1 (see Table 1). The cross-sectional correlations between teachers' self-efficacy and student-rated teaching quality were positive and small for learning support and cognitive activation and were close to zero for classroom management. Only the associations between self-efficacy and learning support were significant (see Table 1).

3.2. Reciprocal relations between teacher motivation and teaching quality: the classical cross-lagged panel approach

The longitudinal relations between teachers' self-efficacy and the three dimensions of student-reported teaching quality were examined with a classical cross-lagged panel approach (see Table 3). As shown in Table 3, the auto-regressive paths across time points (column (5) for the teacher motivation scales; column (6) for the teaching quality scales) are large in many cases (ranging from 0.39 to 0.90) and thus indicate high stability over time, both for the 12 month (Time 1) and the 6 month time (Time 2) interval. None of the three possible cross-lagged effects of teachers' self-efficacy on dimensions of teaching quality (see column (7) in Table 3) were significant across the 6 month interval (Time 2 to Time 3) and the 12 month interval (Time 1 to Time 2). Thus, no longitudinal effects of self-efficacy on teaching quality could be identified. None of the three possible cross-lagged effects of the teaching quality dimensions on teacher self-efficacy (see column (8) in Table 3) was significant for the 6 or the 12 month interval. Longitudinal predictive effects of teaching quality on self-efficacy were therefore not confirmed. Hypothesis 1.1 and 1.2 were not supported in these analyses. A pattern of cross-lagged associations between self-efficacy and teaching quality failed to emerge.

None of the three possible cross-lagged effects of teachers' enthusiasm for teaching on dimensions of teaching quality (see column (7) in Table 3) were significant for the 12 month interval; and two cross-lagged paths were significant for the 6 month interval (positive effects of enthusiasm for teaching on learning support and cognitive activation of 0.18 and 0.26). Thus, for the longer time period, no longitudinal effects of teachers' enthusiasm

Table 4
Random intercepts cross-lagged panel analyses between teacher motivation and teaching quality.

Model	Variables		Bivariate correlations *r* (m, t)			Auto-regressive coefficients		Cross-lagged coefficients		Correlation RI_mRI_t	Loadings λRIm	Loadings λRIt
(1)	(2) m	(3) t	(4) Time 1	(5) Time 2	(6) Time 3	(7) $m_1 \rightarrow m_2$/ $m_2 \rightarrow m_3$	(8) $t_1 \rightarrow t_2$/ $t_2 \rightarrow t_3$	(9) $m_1 \rightarrow t_2$/ $m_2 \rightarrow t_3$	(10) $t_1 \rightarrow m_2$/ $t_2 \rightarrow m_3$	(11)	(12)	(13)
1	Teacher self-efficacy	Classroom management	−0.08	−0.10	−0.51	−0.10/ 0.15*	0.33*/ 0.47***	0.04/−0.07	0.05/0.06	−0.05	0.83−0.87	0.52−0.76
2		Learning support	0.22	0.42*	−0.05	−0.09/0.10	0.08/ 0.54***	−0.02/ −0.07	−0.03/ 0.18*	0.24	0.84−0.87	0.48−0.78
3		Cognitive activation	−0.08	0.14	0.47	−0.12/ 0.13*	0.41*/ 0.79***	0.00/0.12	0.13/−0.02	– [a]	0.85−0.88	– [a]
4	Enthusiasm for teaching	Classroom management	0.00	−0.30	0.24	−0.14/0.18	0.32*/ 0.50***	0.05/0.13	0.02/0.18	0.68**	0.72−0.95	0.53−0.79
5		Learning support	−0.11	−0.40	−0.17	−0.08/0.12	0.17/ 0.56***	−0.15/ .−0.02	−0.01/0.07	0.69**	0.64−0.87	0.47−0.73
6		Cognitive activation	0.15	0.76**	0.22	−0.06/0.15	0.28/0.51*	−0.17/0.30	0.11/0.05	– [a]	0.67−0.83	– [a]

Note. m = motivation (teacher self-efficacy or enthusiasm for teaching); t = teaching quality indicator (student-reported classroom management, learning support or cognitive activation); RI = random intercept. Time 1 to Time 2 = 12 months; Time 2 to Time 3 = 6 months. Column (1) indicates the model number. Columns (2) and (3) indicate which variables were included in each cross-lagged panel analysis. Columns (4), (5) and (6) indicate the correlations between these variables at Time 1 (Column 4), Time 2 (Column 5), and Time 3 (Column 6). Columns (7) and (8) indicate the estimated autoregressive paths for one variable (i.e., its stability) across time points. Column (9) indicates cross-lagged effects of motivation at Time 1/2 on teaching quality at Time 2/3, whereas Column (10) shows the analogous cross-lagged effects of teaching quality at Time 1/2 on motivation at Time 2/3. Column (11) indicates the correlation between the trait factors of the variables. Column (12) indicates the loadings of Time 1 to Time 3 measures of motivation on the motivation trait factor, whereas column (13) shows the analogous loadings of the Time 1 to Time 3 measures of teaching quality on the teaching quality trait factor.
***$p < 0.001$; **$p < 0.01$: *$p < 0.05$. All *p* levels are reported one-tailed.
[a] The variance of the trait factor of cognitive activation was not significant; thus, no loadings and correlations regarding the trait factor are reported.

on teaching quality could be identified, whereas such effects existed for the shorter time period. One of the three possible cross-lagged effects of the teaching quality dimensions on enthusiasm for teaching (see column (8)) was significant for the 12 month interval (a positive effect of classroom management on enthusiasm of 0.24); and one was significant for the 6 month interval (a positive effect of learning support on enthusiasm of 0.32). Longitudinal predictive effects of teaching quality on enthusiasm for teaching thus existed for the longer as well as the shorter time period. Hypothesis 2.1 and 2.2 were only partly supported, since only few significant cross-lagged effects between enthusiasm for teaching and teaching quality emerged.

The analyses provide some support for Hypothesis 3, according to which stronger cross-lagged effects would emerge for the shorter rather than the longer time lag (6 versus 12 months). Only one significant cross-lagged effect was found across the 12-month time interval (a positive effect of classroom management on enthusiasm for teaching), and a total of three cross-lagged effects were found across the 6-month interval (positive effects of enthusiasm on learning support and cognitive activation; and a positive effect of learning support on enthusiasm).

3.3. *Reciprocal relations between teacher motivation and teaching quality: the random intercepts cross-lagged panel approach*

A random intercepts cross-lagged panel analysis was conducted as a next step to estimate stable inter-individual differences in teacher self-efficacy and teaching quality (see Models 1 to 3 in Table 4). The model fit was acceptable across all analyses (Table 5; Schermelleh-Engel, Moosbrugger, & Müller, 2003).

Table 5
Fit indices of the random intercept cross-lagged panel models.

Model	χ^2	*df*	CFI	RMSEA	$SRMR_{within}$	$SRMR_{between}$
1	1.22	1	1.00	0.01	0.00	0.01
2	0.56	1	1.00	0.00	0.00	0.01
3	4.80	3	1.00	0.01	0.00	0.09
4	0.21	1	1.00	0.00	0.00	0.01
5	4.54	1	0.99	0.03	0.00	0.07
6	10.99	1	0.97	0.05	0.00	0.09

Note. The model numbers refer to Table 4.

The loadings on the random intercepts indicate high to very high stabilities for the two teacher motivation characteristics and moderate to high stabilities for the teaching quality dimensions (see Table 4, columns (12) and (13)). However, there was no significant trait variance for cognitive activation, which varied greatly across time points. The respective association could therefore not be computed in a meaningful way. The associations between the random intercepts of self-efficacy and the teaching quality dimensions were non-existing to weak (max. of $r = 0.24$; see column (11) in Table 4). Thus, the trait aspects of self-efficacy and teaching quality were only weakly or not at all related.

The auto-regressive paths across time points, and thus the individual carry-over effects (i.e., effects on the individual level that persist across measurement points), are mostly not significant for self-efficacy (-0.12 to 0.15; see column (7) in Table 4), and are mostly medium to large for all teaching quality dimensions (0.08−0.79; see column (8) in Table 4).

Controlling for trait-like stability in teacher motivation and teaching quality, none of the six possible cross-lagged effects of teachers' self-efficacy on the three dimensions of teaching quality was significant (see column (9)); one of the six possible cross-lagged effects of the teaching quality dimensions on self-efficacy was significant (a positive effect of learning support on self-efficacy, see column (10)). Satorra-Bentler scaling-corrected chi-square difference tests (one-tailed) revealed, however, that this significant cross-lagged path from Time 2 learning support to Time 3 self-efficacy was not significantly different from the corresponding non-significant path from Time 2 self-efficacy to Time 3 learning support ($\Delta\chi^2_{diff} = 1.96$, $df = 1$, $p = 0.08$). Therefore, we cannot assume that the link from learning support to self-efficacy differs from the reverse effect. The results of the random intercept cross-lagged panel analyses of the associations between teacher enthusiasm for teaching and the three teaching quality

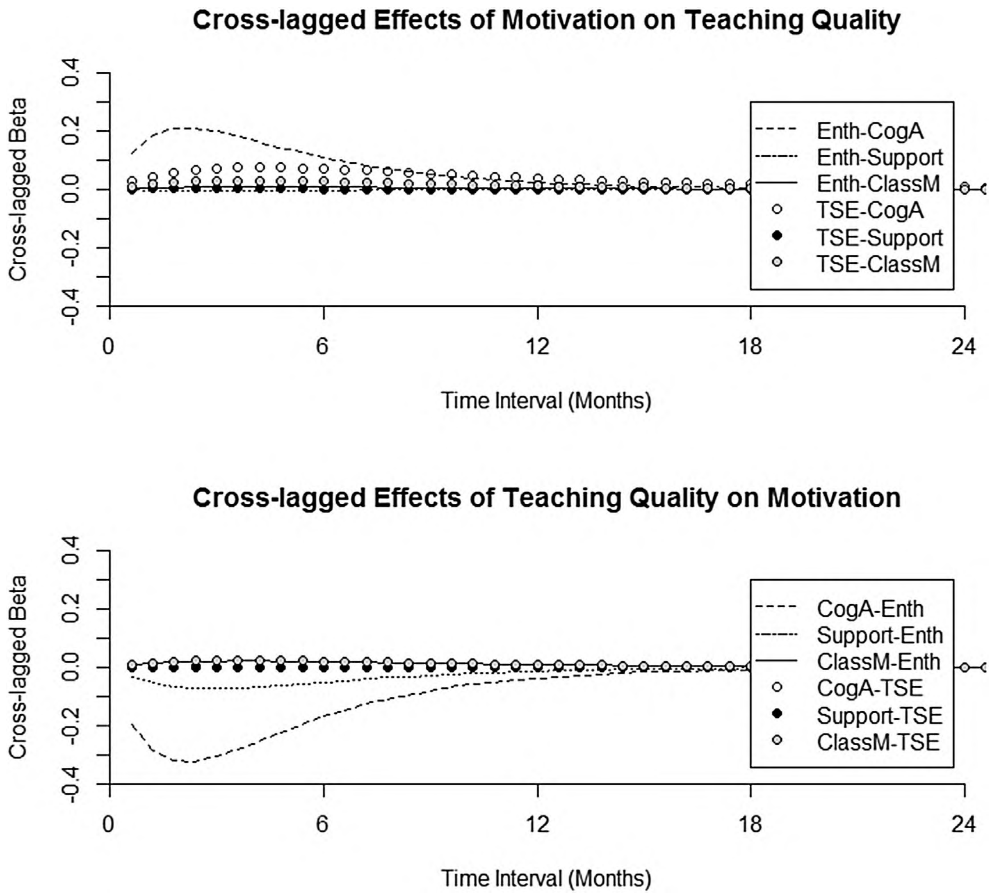

Fig. 3. Estimation of the optimal time lags for investigating the reciprocal effects of teacher motivation and teaching quality within the random intercept cross-lagged panel approach. Enth = enthusiasm; TSE = self-efficacy; CogA = cognitive activation; Support = learning support; ClassM = classroom management.

dimensions are shown in Table 4 (see Models 4 to 6). The model fit was again satisfactory for all analyses (see Table 5). The correlations between the random intercepts of teacher enthusiasm and the teaching quality dimensions ranged from 0.68 to 0.69 (see column (11) in Table 4; again with the exception of cognitive activation), indicating that the trait-like parts of enthusiasm and teaching quality were mostly highly correlated.

The auto-regressive paths across time points, and thus the individual carry-over effects, are not significant for enthusiasm (see column (7) in Table 4), but have mostly medium to large values for all teaching quality dimensions (0.17–0.56; see column (8) in Table 4).

None of the tested cross-lagged effects between teachers' enthusiasm for teaching and the teaching quality dimensions were significant, once trait-like stability in teacher motivation and teaching quality was taken into account (see columns (9) and (10) in Table 4). Thus, longitudinal effects of enthusiasm for teaching on teaching quality or the other way around could not be confirmed.

Consistent with Hypothesis 4, these analyses suggest that almost all cross-lagged associations between teacher motivation and teaching quality revealed with our classical cross-lagged models could be explained with trait-like associations.

3.4. *Identifying optimal time lags for investigating cross-lagged effects*

Analyses of the optimal time lag for the presented cross-lagged models were conducted using an approach described by Dormann and Griffin (2015). The Time 1 and Time 2 data from the random intercepts latent cross-lagged model were used for these analyses. The results are illustrated in Fig. 3. The optimal time lag for analyses focusing on the cross-lagged effects of enthusiasm for teaching would be about 3 months for learning support and classroom management and about 2 months for cognitive activation. The optimal time lag for analyses focusing on the cross-lagged effects of teacher self-efficacy would be about 4 months for learning support and cognitive activation, and about 3 months for classroom management. These findings are generally consistent with our expectations that cross-lagged effects are more likely to occur with shorter time intervals. However, as shown in Fig. 3, even at their expected maximum value, the estimated cross-lagged effects between the teacher motivation variables and the teaching quality dimensions are very small (with the exception of cognitive activation and enthusiasm), indicating no substantial cross-lagged associations independent of the chosen time lag.

4. Discussion

Teacher motivation is often assumed to be an antecedent of desirable teaching behaviors (see e.g. Richardson & Watt, 2010) and thus of teaching quality (e.g., Kunter et al., 2011, 2013). However, motivational characteristics not only have an effect on teachers' teaching quality but can also be influenced by it as has been shown initially by Holzberger et al. (2013). Additionally, relations between motivational characteristics and teaching quality could as well be due to third variables. The present study aimed at answering the question of how associations between teaching-related motivations and student-reported teaching quality are shaped longitudinally in a sample of secondary-level math teachers and their students across three time points.

Our analyses using classical cross-lagged panel models showed some of the expected cross-lagged effects of teacher motivation on teaching quality and vice versa. However, controlling for stable inter-individual differences (see Hamaker et al., 2015), we found no systematic evidence in support of cross-lagged effects. The loadings of the state measures for teacher motivation and teaching quality on the trait factors (random intercepts for teacher motivation and teaching quality) indicate that all investigated characteristics besides cognitive activation were fairly to highly stable so that reciprocal associations between teaching quality and teacher motivation are not likely to occur, even over a period of one and a half years and with three measurement points. An optimal time lag analysis indicated that a shorter time frame is more likely to reveal potential cross-lagged effects, but that these effects are likely to be weak regardless of the time frame with only one exception (the longitudinal relation between cognitive activation and enthusiasm for teaching). The results instead indicate that teachers whose students consistently perceive higher levels of teaching quality tended also to consistently report higher enthusiasm, but that these relations cannot be explained using a randomly chosen time span in the working life of a teacher. The stable part in teachers' self-efficacy, on the contrary, was relatively independent from the respective stable parts in teaching quality. This high level of stability is consistent with earlier evidence using a similar measure of teacher self-efficacy (Holzberger et al., 2013).

4.1. The relation between enthusiasm and teaching quality: explanations and implications

Over and above potential longitudinal effects of enthusiasm for teaching on teaching quality and vice versa, third variables were mentioned as a possible explanation for existing relations between enthusiasm and teaching quality. The high correlations of the trait parts of enthusiasm and teaching quality indicate that third variables indeed could have a considerable influence on both, enthusiasm and teaching quality. Such third variables might include teacher, class, and school characteristics. Teacher variables such as teachers' professional knowledge or their beliefs could shape their enthusiasm on the one hand and their teaching on the other hand. Correlational patterns in the study of Kunter et al. (2013) indicate, however, that significant relations between pedagogical content knowledge and enthusiasm as well as teaching quality do not exist consistently. The same is true for constructivist beliefs and enthusiasm as well as teaching quality. Class characteristics are another set of possible third variables. Kunter et al. (2011) could show that enthusiasm is related to class characteristics such as class size and average student achievement. Additionally, teaching enjoyment (which is closely related to the concept of enthusiasm for teaching, see section 1.1) has been shown to vary significantly between classes (Frenzel, Becker-Kurz, Pekrun, & Goetz, 2015). Teaching quality, too, seems to differ between classes a teacher teaches (Kokkinou & Kyriakides, 2016). To what degree class characteristics shape enthusiasm und teaching simultaneously, however, remains an open question for future research. School characteristics (e.g., school leadership or teachers' organizational commitment), finally, are another set of possible third variables which have not been investigated so far.

4.2. The relation between self-efficacy and teaching quality: explanations and implications

Whereas third variables might be an explanation for existing relations between enthusiasm and teaching quality, this is not the case for self-efficacy as neither longitudinal relations nor relations between their trait aspects could be found. Alternative explanations are research focus, matching issues, and samples.

Most investigations of teacher motivation have exclusively focused on relations to aspects of teaching quality on the teacher/class level. There is, however, evidence that teachers' motivations can be shaped by their relationships with single students even more strongly than by their relationships with a whole class (Lortie, 1975). If the whole class is not the reference norm for building motivation and perhaps also not for the influence of motivation on improving instruction (i.e., improving instruction not for all students but rather for some of them), then future research would need to take a closer look at the variation within classrooms. Qualitative approaches might be particularly useful for answering this question.[3]

Furthermore, self-efficacy and teaching quality might not be related in the current study because they do not refer to the same entities (cf. Bandura, 2012; Wheatley, 2005). Whereas enthusiasm was assessed in a context-specific way, focusing on the class in which student-reported teaching quality was measured, self-efficacy was assessed using the original version of the teacher self-efficacy scale. The scale refers to teaching in general and not to a specific class which is common in teacher self-efficacy research (e.g., Tschannen-Moran & Woolfolk Hoy, 2001). Holzberger et al. (2013) modified the scale to refer to a specific class, which may contribute to an increased domain-specificity and thus to stronger associations with class-specific teaching quality. Assessing self-efficacy with respect to a specific target class thus seems reasonable for future studies, especially because it could be shown that self-efficacy can vary across different classes of the same teacher (see, e.g., Raudenbush, Rowan, & Cheong, 1992).

Teachers' self-efficacy also depends on their career stage: For early-career and late career teachers, substantial changes in self-efficacy can be expected (Klassen & Chiu, 2010; see alsoWoolfolk Hoy and Burke Spero, 2005). Consequently, it may be more likely to reveal longitudinal influences among early-career and late-career teachers relative to mid-career teachers. A first hint that relations between teacher motivation and teaching quality indeed differ for different career stages, can be gained by comparing the study of Holzberger et al. (2013) and our study. For the sample of Holzberger et al. – consisting to a considerable degree of late-stage career teachers (*M* job experience = 22 years) – effects of student-perceived teaching quality on self-efficacy could be found. In the present study – covering mainly mid-career teachers (*M* job experience = 13 years) – no longitudinal relations between self-efficacy and teaching quality could be identified. If teachers' self-efficacy is indeed less stable during certain stages of a teacher's career but very stable during others, theories on teachers' self-efficacy need to be further developed so that career stages are explicitly included. For teacher training and professional development, these considerations indicate that it might be especially useful to support teachers with respect to their self-efficacy in their early and late career stages. If self-efficacy and teaching quality are

[3] In accordance with the most common approach in cognitive survey research (e.g., Willis, 2005), one could let teachers think aloud (Ericsson & Simon, 1993), while answering quantitative scales on their motivation as well as conducting interviews with verbal prompts afterwards to see whether teachers refer to only some students when reflecting on their motivation.

closely intertwined during these career stages, it seems also very promising to not only foster self-efficacy, but at the same time teaching quality.

Furthermore, further research is needed to shed more light on the mechanisms linking teacher motivation and teaching quality. First, a sound theoretical model is required to explicate these mechanisms. Part of such a model could, for example, be the assumption that high self-efficacy lets teachers perceive a need to focus on the teaching job and to continuously improve teaching which, in turn, leads to corresponding actions by the teacher (e.g., effort towards teaching). Hypotheses which are contrary to these assumptions should, however, also be taken into account. According to Wheatley (2005) it might also be the case that lacking self-efficacy may be more likely than high self-efficacy to lead to a perceived need to and corresponding actions aiming towards personal improvement. Second, in bringing these assumptions to an empirical test, we need to think carefully about the time intervals that, from a theoretical perspective, make it most likely to uncover the assumed mechanisms. Third, for detecting these mechanisms, it seems worthwhile to also use more fine-grained methods (e.g., daily logs, Borko et al., 2007, or experience sampling methods, Csikszentmihalyi & Larson, 2014) or more direct approaches such as experimental and quasi-experimental studies (e.g., see Tschannen-Moran & McMaster, 2009) to unfold what is actually happening with respect to the relation between teacher motivation and teaching quality.

4.3. Limitations and further directions

Our study was based on a sample of teachers who were investigated longitudinally over the course of 18 months. As it was necessary for answering our research questions to restrict the longitudinal sample to those teachers who were teaching mathematics in the same class in two subsequent years, substantial parts of the sample were not investigated at Times 2 and 3. Our analyses suggest, however, that teachers who were retained in the sample did not differ systematically from those who had to be excluded.

Another limitation refers to the fact that the analyses were not conducted on a latent level due to an insufficient sample of teachers. Analyses of latent constructs may have produced stronger results. However, since the size of the regression coefficients was small (independent of their standard errors) and the loadings on the trait factors in the random-intercept models were large, results are not expected to be substantially different if we were using a latent approach.

Estimating optimal time lags is highly relevant for future research. However, whether the estimated lags are accurate, depends on several assumptions. An aspect that we were not able to control due to our use of observed variables is measurement error, which could lead to biased estimates of optimal time lags. Thus, we should be rather careful to not over-interpret single estimates in the optimal time lag analyses. At the same time, we must point out that the very high stability of teacher motivation constructs—at least for self-efficacy—is consistent with earlier research using latent variables (Holzberger et al., 2013, $r_{tt} = 0.85$). Accordingly, our estimation of small reciprocal effects, regardless of the time lag, is plausible. Another implicit assumption of optimal time lag analyses is that psychological mechanisms are the same for different time lags. To what extent this is justified needs to be investigated in future studies. This is not only important to test the trustworthiness of optimal time lag analyses but also to check whether prior studies with different time lags can be compared at all.

Finally, we used student reports as an indicator of teaching quality. Student ratings, aggregated on the class level, have many advantages compared to teacher self-ratings (Clausen, 2002; Kunter & Baumert, 2006; Lüdtke, Robitzsch, Trautwein, & Kunter, 2009). However, in order to be suitable for a representation of shared perceptions of students in a class, students' ratings need to capture sufficient variability on the class/teacher level compared to the individual student level, and to have sufficient reliability not only on the individual but also on the class level. In the present study, classroom management emerged as the dimension with the highest levels of agreement among students. This is plausible, because classroom management ratings typically require relatively low levels of inference and instead rely on rather observable indicators such as class disruptions or time on task. Cognitive activation, in contrast, more strongly depends on students' idiosyncratic perceptions. Accordingly, lower levels of reliability on the class level have been documented, both in the present study (especially at Time 1) and in prior research (see e.g., Kunter et al., 2008). Assessments of cognitive activation might be most suitable as a student-level indicator of individual perceptions of instruction, and alternative measures might be necessary for capturing this dimension on the classroom level (e.g., observations by trained external evaluators). One option to improve student ratings might be the time of investigation within the school year as the amount of shared student perceptions in a class increased over time. This indicates that it might be useful to measure teaching quality not too early in the school year when student perceptions are used to measure teaching quality.

4.4. Conclusions

Although it is often taken for granted that teacher motivation has an influence on teaching quality, the present study showed that longitudinal effects of teacher motivation on teaching quality or vice versa do not necessarily exist as both were rather stable over time in the present sample. These stable parts, however, were substantially related, at least for teachers' enthusiasm and teaching quality. It seems to be promising to focus on early- and late-stage career teachers in further research, because reciprocal influences might exist to a larger extent for these groups.

Acknowledgements

This research was funded by a grant from the German Federal Ministry of Education and Research to Oliver Dickhäuser (01 HJ 0901) and Markus Dresel (01 HJ 0902). We thank Christian Dormann for helpful comments on the application of his formula for optimal time lags.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.tate.2017.03.023.

References

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: Issues and techniques* (pp. 243–277). Mahwah, NJ: Lawrence Erlbaum Associates.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: Freeman.

Bandura, A. (2012). On the functional properties of perceived self-efficacy revisited. *Journal of Management, 38*, 9–44. http://dx.doi.org/10.1177/0149206311410606.

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.

Borko, H., Kuffner, K. L., Arnold, S. C., Creighton, L., Stecher, B., Martinez, F., et al. (2007). *Using artifacts to describe instruction: Lessons learned from studying reform-oriented instruction in middle school mathematics and science*. CSE Technical Report 705. Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing.

Caprara, G. V., Barbaranelli, C., Borgogni, L., & Steca, P. (2003). Efficacy beliefs as determinants of teachers' job satisfaction. *Journal of Educational Psychology, 95*, 821–832. http://dx.doi.org/10.1016/j.jsp.2006.09.001.

Clausen, M. (2002). *Qualität von Unterricht – Eine Frage der Perspektive?* [*Quality of instruction – A question of perspective?*] Munster: Waxmann.

Cohen, D. K. (1993). *Teaching for understanding: Challenges for policy and practice*. San Francisco: Jossey-Bass.

Cornelius-White, J. (2007). Learner-centered teacher-student relationships are effective: A meta-analysis. *Review of Educational Research, 77*, 113–143.

Csikszentmihalyi, M., & Larson, R. (2014). Validity and reliability of the experience-sampling method. In M. Csikszentmihalyi (Ed.), *Flow and the foundations of positive psychology*. Dordrecht: Springer. pp. 35–54.

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.

Dewey, John (1916). *Democracy and education: An introduction to the philosophy of education*. New York, NY: Macmillan.

Dicke, T., Parker, P. D., Holzberger, D., Kunter, M., & Leutner, D. (2015). Beginning teachers' efficacy and emotional exhaustion: Latent changes, reciprocity, and the influence of professional knowledge. *Contemporary Educational Psychology, 41*, 62–72. http://dx.doi.org/10.1016/j.cedpsych.2014.11.003.

Dormann, C., & Griffin, M. (2015). Optimal time lags in panel studies. *Psychological Methods, 20*, 489–505.

Eccles, Y. (2009). Who am I and what am I going to do with my life? Personal and collective identities as motivators of action. *Educational Psychologist, 44*, 78–89. http://dx.doi.org/10.1080/00461520902832368.

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology, 53*, 109–132. http://dx.doi.org/10.1146/annurev.psych.53.100901.135153.

Enders, C. K., Mistler, S. A., & Keller, B. T. (2016). Multilevel multiple imputation: A review and evaluation of joint modeling and chained equations imputation. *Psychological Methods, 21*, 222–240. http://dx.doi.org/10.1037/met0000063.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press.

Fauth, B., Decristan, J., Rieser, S., Klieme, E., & Büttner, G. (2014). Student ratings of teaching quality in primary school: Dimensions and prediction of student outcomes. *Learning and Instruction, 29*, 1–9. http://dx.doi.org/10.1016/j.learninstruc.2013.07.001.

Frenzel, A. C.*, Becker-Kurz, B.*, Pekrun, R., & Goetz, T. (2015). Teaching this class drives me nuts! – Examining the person and context specificity of teacher emotions. *PLoS One, 10*. http://dx.doi.org/10.1371/journal.pone.0129630 (* authors contributed equally).

Frenzel, A. C., Goetz, T., Lüdtke, O., Pekrun, R., & Sutton, R. E. (2009). Emotional transmission in the classroom: Exploring the relationship between teacher and student enjoyment. *Journal of Educational Psychology, 101*, 705–716. http://dx.doi.org/10.1037/a0014695.

Graham, S., & Weiner, B. (1996). Theories and principles of motivation. In D. C. Berliner, & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 63–84). New York, NY: Macmillan.

Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. (2015). A critique of the cross-lagged panel model. *Psychological Methods, 20*, 102–116.

Hattie, J. (2009). *Visible learning: A synthesis of meta-analyses relating to achievement*. London: Routledge.

Hein, V., Ries, F., Pires, F., Caune, A., Emeljanovas, A., Heszteráné Ekler, J. H., et al. (2012). The relationship between teaching styles and motivation to teach among physical education teachers. *Journal of Sports Science and Medicine, 11*, 123–130.

Holzberger, D., Phillip, A., & Kunter, M. (2016). Ein Blick in die Black-Box: Wie der Zusammenhang von Unterrichtsenthusiasmus und Unterrichtshandeln von angehenden Lehrkräften erklärt werden kann. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 48*, 90–105. http://dx.doi.org/10.1026/0049-8637/a000150.

Holzberger, D., Philipp, A., & Kunter, M. (2013). How teachers' self-efficacy is related to teaching quality: A longitudinal analysis. *Journal of Educational Psychology, 105*, 774–786. http://dx.doi.org/10.1037/a0032198.

Holzberger, D., Philipp, A., & Kunter, M. (2014). Predicting teachers' instructional behaviors: The interplay between self-efficacy and intrinsic needs. *Contemporary Educational Psychology, 39*(2), 100–111. http://dx.doi.org/10.1016/j.cedpsych.2014.02.001.

Jamil, F. M., Downer, J. T., & Pianta, R. C. (2012). Association of pre-service teachers' performance, personality, and beliefs with teacher self-efficacy at program completion. *Teacher Education Quarterly, 39*, 119–138.

Klassen, R. M., & Chiu, M. M. (2010). Effects on teachers' self-efficacy and job satisfaction: Teacher gender, years of experience, and job stress. *Journal of Educational Psychology, 102*, 741–756. http://dx.doi.org/10.1037/a0019237.

Klassen, R. M., & Tze, V. M. C. (2014). Teachers' self-efficacy, personality, and teaching effectiveness: A meta-analysis. *Educational Research Review, 12*, 59–76. http://dx.doi.org/10.1016/j.edurev.2014.06.001.

Klieme, E., & Rakoczy, K. (2003). Unterrichtsqualität aus Schülerperspektive: Kulturspezifische Profile, regionale Unterschiede und Zusammenhänge mit Effekten von Unterricht. In J. Baumert, C. Artelt, E. Klieme, J. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, & K.-J. Tillmann (Eds.), PISA 2000. Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland (pp. 334–359) [PISA 2000. A differentiated view on the German Laenders]. Opladen: Leske + Budrich.

Klassen, R. M., Tze, V. M. C., Betts, S. M., & Gordon, K. A. (2011). Teacher efficacy research 1998–2009: Signs of progress or unfulfilled promise? *Educational Psychological Review, 23*, 21–43. http://dx.doi.org/10.1007/s10648-010-9141-8.

Kokkinou, E., & Kyriakides, L. (2016). Assessing teacher effectiveness: Investigating the consistency of teacher behavior in different classrooms. In *Paper presented at the EARLI SIGs 18 and 23 joint conference in Oslo*.

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

Kunter, M. (2005). *Multiple Ziele im Mathematikunterricht* [*Multiple goals in mathematics instructions*]. Münster, Germany: Waxmann.

Kunter, M. (2011). Forschung zur Lehrermotivation. In E. Terhart, H. Bennewitz, & M. Rothland (Eds.), Handbuch der Forschung zum Lehrerberuf (pp. 527–539) [Handbook of research on the teacher job]. Münster, Germany: Waxmann.

Kunter, M., & Baumert, J. (2006). Who is the expert? Construct and criteria validity of student and teacher ratings of instruction. *Learning Environments Research, 9*, 231–251. http://dx.doi.org/10.1007/s10984-006-9015-7.

Kunter, M., Baumert, J., Voss, T., Klusmann, U., Richter, D., & Hachfeld, A. (2013). Professional competence of teachers: Effects on instructional quality and student development. *Journal of Educational Psychology, 10*, 805–820. http://dx.doi.org/10.1037/a0032583.

Kunter, M., Frenzel, A., Nagy, G., Baumert, J., & Pekrun, R. (2011). Teacher enthusiasm: Dimensionality and context specificity. *Contemporary Educational Psychology, 36*, 289–301. http://dx.doi.org/10.1016/j.cedpsych.2011.07.001.

Kunter, M., & Holzberger, D. (2014). Loving teaching: Research on teachers' intrinsic orientations. In P. W. Richardson, S. A. Karabenick, & H. M. G. Watt (Eds.), *Teacher motivation: Theory and practice* (pp. 83–99). New York, NY: Routledge.

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. http://dx.doi.org/10.1016/j.learninstruc.2008.06.008.

LeBreton, J. M., & Senter, J. L. (2008). Answers to 20 questions about interrater reliability and interrater agreement. *Organizational Research Methods, 11*, 815–852.

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. http://dx.doi.org/10.1016/j.learninstruc.2008.11.001.

Lohman, M. C. (2006). Factors influencing teachers' engagement in informal learning activities. *Journal of Workplace Learning, 18*, 141–156. http://dx.doi.org/10.1108/13665620610654577.

Lortie, D. C. (1975). *Schoolteacher: A sociological study*. Chicago: University of Chicago Press.

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. http://dx.doi.org/10.1016/j.cedpsych.2008.12.001.

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. http://dx.doi.org/10.1037/0003-066X.59.1.14.

McArdle, J. J. (2009). Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology, 60*, 577–605. http://dx.doi.org/10.1146/annurev.psych.60.110707.163612.

Morris-Rothschild, B., & Brassard, M. R. (2006). Teachers' conflict managment styles: The role of attachment styles and classroom managment efficacy. *Journal of School Psychology, 44*, 105–121.

Muthén, L. K., & Muthén, B. O. (1998–2012). *Mplus users' guide* (seventh ed.). Los Angeles, CA: Muthén & Muthén.

Patrick, H., Turner, J. C., Meyer, D. K., & Midgley, C. (2003). How teachers establish psychological environments during the first days of school: Associations with avoidance in mathematics. *Teachers College Record, 105*, 1521–1558.

Pelletier, L. G., Séguin-Lévesque, C., & Legault, L. (2002). Pressure from above and pressure from below as determinants of teachers' motivation and teaching behaviors. *Journal of Educational Psychology, 94*, 186–196. http://dx.doi.org/10.1037//0022-0663.94.1.186.

Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher, 38*, 109–119. http://dx.doi.org/10.3102/0013189X09332374.

Praetorius, A.-K., Pauli, C., Reusser, K., Rakoczy, K., & Klieme, E. (2014). One lesson is all you need? Stability of instructional quality across lessons. *Learning and Instruction, 31*, 2–12. http://dx.doi.org/10.1016/j.learninstruc.2013.12.002.

Rakoczy, K. (2008). *Motivationsunterstützung im Mathematikunterricht: Unterricht aus der Perspektive von Lernenden und Beobachtern* [*Motivational support in mathematics lessons: Instruction from the perspectives of learners and observers*]. Münster, Germany: Waxmann.

Ramm, G., Prenzel, M., Baumert, J., Blum, W., Lehmann, R., Leutner, D., et al. (2006). *PISA 2003 – Dokumentation der Erhebungsinstrumente* [*PISA 2003 – Documentation of the survey questionnaires*]. Münster, Germany: Waxmann.

Raudenbush, S. W., Rowan, B., & Cheong, Y. F. (1992). Contextual effects on the self-perceived efficacy of high school teachers. *Sociology of Education, 65*, 150–167. http://dx.doi.org/10.2307/2112680.

Reyes, M. R., Brackett, M. A., Rivers, S. E., White, M., & Salovey, P. (2012). Classroom emotional climate, student engagement, and academic achievement. *Journal of Educational Psychology, 104*, 700–712. http://dx.doi.org/10.1037/a0027268.

Richardson, P. W., & Watt, H. M. G. (2010). Current and future directions in teacher motivation research. In T. C. Urdan, & S. A. Karabenick (Eds.), *The decade ahead: Applications and contexts of motivation and achievement: 16A. Advances in motivation and achievement* (pp. 139–173). Bingley, UK: Emerald Publishing Group.

Ross, J. A., & Bruce, C. D. (2007). Teacher self-assessment: A mechanism for facilitating professional growth. *Teaching and Teacher Education, 23*, 146–159.

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. http://dx.doi.org/10.1037/0022-0663.99.4.761.

Ryan, M. R., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 51*, 68–78.

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *MPR-online, 8*, 23–74.

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. http://dx.doi.org/10.1024//1010-0652.14.1.12.

Schwarzer, R., & Schmitz, G. S. (1999). Skala Lehrer-Selbstwirksamkeitserwartung. In R. Schwarzer, & M. Jerusalem (Eds.), *Skalen zur Erfassung von Lehrer- und Schülermerkmalen* (pp. 60–61). Berlin, Germany: Freie Universität Berlin [Scales for measuring teacher and student characteristics].

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

Skaalvik, E. M., & Skaalvik, S. (2007). Dimensions of teacher self-efficacy and relations with strain factors, perceived collective teacher efficacy, and teacher burnout. *Journal of Educational Psychology, 99*, 611–625. http://dx.doi.org/10.1037/0022-0663.99.3.611.

Soodak, L. C., & Podell, D. M. (1998). Teacher efficacy and the vulnerability of the difficult-to-teach student. In J. Brophy (Ed.), *Advances in research on teaching* (Vol. 7, pp. 75–109). Greenwich, CT: JAI Press.

Tschannen-Moran, M., & Woolfolk Hoy, A. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education, 17*, 783–805. http://dx.doi.org/10.1016/S0742-051X(01)00036-1.

Tschannen-Moran, M., & McMaster, P. (2009). Sources of self-efficacy: Four professional development formats and their relationship to self-efficacy and implementation of a new teaching strategy. *The Elementary School Journal, 110*, 228–245. http://dx.doi.org/10.1086/605771.

Tschannen-Moran, M., Woolfolk Hoy, A., & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research, 68*, 202–248. http://dx.doi.org/10.3102/00346543068002202.

Vieluf, S., Kunter, M., & van de Vijver, F. J. R. (2013). Teacher self-efficacy in cross-national perspective. *Teaching and Teacher Education, 35*, 92–103.

Voelkle, M. C., Oud, J. H. L., Davidov, E., & Schmidt, P. (2012). A SEM approach to continuous time modeling of panel data: Relating authoritarism and anomia. *Psychological Methods, 17*, 176–192. http://dx.doi.org/10.1037/a0027543.

Ware, H., & Kitsantas, A. (2007). Teacher and collective efficacy beliefs as predictors of professional commitment. *Journal of Educational Research, 100*, 303–310.

Wheatley, K. F. (2005). The case for reconceptualizing teacher efficacy research. *Teaching and Teacher Education, 21*, 747–766. http://dx.doi.org/10.1016/j.tate.2005.05.009.

Wigfield, A., & Eccles, J. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology, 25*, 68–81. http://dx.doi.org/10.1006/ceps.1999.1015.

Wild, E. (1999). *Elterliche Erziehung und schulische Lernmotivation* [*Parental education and academic motivation*] (Unpublished post-doctoral thesis). Germany: University of Mannheim.

Williams, L. J., Hartman, N., & Cavazotte, F. (2010). Method variance and marker variables: A review and comprehensive CFA marker technique technique. *Organizational Research Methods, 13*, 477–514. http://dx.doi.org/10.1177/1094428110366036.

Willis, G. (2005). *Cognitive interviewing: A tool for improving questionnaire design*. Thousand Oaks, CA: Sage.

Woolfolk Hoy, A. (2008). What motivates teachers? Important work on a complex question. *Learning and Instruction, 18*, 492–498. http://dx.doi.org/10.1016/j.learninstruc.2008.06.007.

Woolfolk Hoy, A., & Burke Spero, R. (2005). Changes in teacher efficacy during the early years of teaching: A comparison of four measures. *Teaching and Teacher Education, 21*, 343–356. http://dx.doi.org/10.1016/j.tate.2005.01.007.

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. http://dx.doi.org/10.3102/0034654315626801.