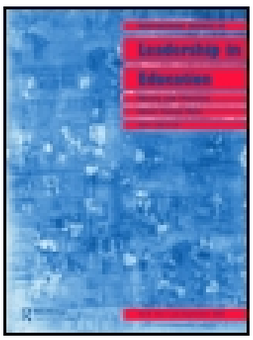

International Journal of Leadership in Education
Theory and Practice

ISSN: 1360-3124 (Print) 1464-5092 (Online) Journal homepage: https://www.tandfonline.com/loi/tedl20

Exploring improvement in teachers' instructional support: classifying and analyzing patterns of change in a national initiative on classroom management

Sigrun K. Ertesvåg

To cite this article: Sigrun K. Ertesvåg (2019): Exploring improvement in teachers' instructional support: classifying and analyzing patterns of change in a national initiative on classroom management, International Journal of Leadership in Education, DOI: 10.1080/13603124.2019.1613567

To link to this article: https://doi.org/10.1080/13603124.2019.1613567

© 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

Published online: 20 May 2019.

Submit your article to this journal

Article views: 1086

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=tedl20

INTERNATIONAL JOURNAL OF LEADERSHIP IN EDUCATION
https://doi.org/10.1080/13603124.2019.1613567

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

Exploring improvement in teachers' instructional support: classifying and analyzing patterns of change in a national initiative on classroom management

Sigrun K. Ertesvåg

Norwegian Centre for Learning Environment and Behavioural Research in Education, University of Stavanger, Stavanger, Norway

ABSTRACT

What teachers do in the classroom is a key determinant of student learning processes and outcomes. The current study explores the levels of and improvement in teachers' instructional support throughout a school-wide initiative that aimed to develop classroom management in Norwegian lower secondary schools. The sample contained 227 teachers from nine schools. The results of growth mixture modeling (GMM) indicate that the majority of teachers did not improve. The teachers who did improve reported higher initial instructional support levels than the non-improving teachers. The improving teachers also reported higher job satisfaction and less stress from student behavior. No significant differences were found between the two groups with respect to emotional exhaustion and stress from workload. The study illustrates an approach relevant for evaluating improvement in teachers' practice within specific interventions as well as in ongoing professional development. The results may have implications for the length, focus and content of professional development for teachers. The study provides knowledge that may be relevant to school and district leaders initiating improvement efforts in their school(s).

Introduction

Researchers and practitioners have long known that what teachers do in the classroom is a key determinant of student learning outcomes (e.g., Kane & Staiger, 2012; Kersting, Givvin, Thompson, Santagata, & Stigler, 2012; Klieme, Pauli, & Reusser, 2009; Pianta, Hamre, & Allen, 2012). In this context, a number of descriptions of teacher–student interaction have been put forth in the educational and developmental literature (e.g., Fauth, Decristan, Rieser, Klieme, & Büttner, 2014; Kunter & Baumert, 2006; Pianta et al., 2012; Wagner, Göllner, Helmke, Trautwein, & Lüdtke, 2013; Wubbels & Brekelmans, 2005). Teacher–student interaction has been described in terms of teachers' emotional support, classroom organization and instructional support (Hamre et al., 2013; Pianta, 2016; Pianta et al., 2012). The construct is understood and operationalized differently across the field, but its multidimensionality is found in

CONTACT Sigrun K. Ertesvåg sigrun.ertesvag@uis.no Norwegian Center for Learning Environment and Behavioral Research in Education, University of Stavanger, 4036 Stavanger, Norway

© 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

research in both Europe (e.g., Ertesvåg & Havik, 2018; Fauth et al., 2014; Kunter & Baumert, 2006) and the United States (Hamre et al., 2013; Kane & Staiger, 2012).

Teachers perceive themselves as skilled in emotional support and classroom organization (Ertesvåg, 2011; Gitomer et al., 2014), but little is known about how they perceive their instructional support and improvement in these skills throughout an intervention.

Instructional support comprises teachers' demonstration of consistent, process-oriented feedback; their focus on higher-order thinking skills; their ability to challenge students cognitively; their presentation of new content within a broader, meaningful context; and their application of knowledge in a problem-solving context. It also refers to the strategies that teachers use to enhance student engagement in instructional content (Fauth et al., 2014; Hamre et al., 2013; Klieme et al., 2009; Pianta et al., 2012).

Recent research indicates that the quality of teacher–student interaction is related to teachers' job satisfaction (McCarthy, Lineback, & Reiser, 2015; Virtanen, Vaaland, & Ertesvåg, 2019). Job satisfaction, emotional exhaustion and stress may influence teachers' ability and willingness to take part in professional development (PD) (Hargreaves, 2005; Nir & Bogler, 2008).

Accordingly, the main aim of this study is to explore teachers' perceived improvement in instructional support, which is considered a key aspect of classroom management (Pianta, 2006), throughout a national initiative that aimed to improve lower secondary schools (Grades 8–10, age 13–15 years). The study also aims to explore differences in job satisfaction, emotional exhaustion and stress among teachers who differ in their improvement in instructional support. Moreover, the study illustrates an approach for evaluating improvement in teachers' practice both within specific interventions and in ongoing PD that is not necessarily part of a set intervention. Accordingly, the study may provide knowledge important to school and district leaders initiating improvement efforts in the field.

Instructional support

In the current study, the understanding of instructional support aligns with the Teaching Through Interaction (TTI) framework developed by Robert Pianta, Bridget Hamre and colleagues (e.g., Hamre et al., 2013; Pianta et al., 2012). This framework organizes teacher-student classroom interaction into three broad domains of support for students' development and learning: emotional support, classroom organization and instructional support (Hamre et al., 2013; Pianta et al., 2012). Previous studies have linked these types of classroom interactions to students' academic learning (e.g., Allen et al., 2013; Pianta et al., 2012 for an overview).

Examples of instructional support in line with the TTI framework are teachers‘ use of feedback to expand and extend learning and understanding and to encourage student participation. Another example is the depth of lesson content and the approaches used to help students comprehend the framework, key ideas, and procedures in an academic discipline. At a high level, instructional support refers to the interactions among the teacher and students that lead to an integrated understanding of facts, skills, concepts and principles. A third example is the teachers' facilitation of teaching so that the students are engaged in higher-level thinking skills through the application of knowledge and skills to novel and/or open-ended problems, tasks, and questions.

Opportunities for engaging in metacognition, i.e., thinking about thinking, are also included (Pianta, Hamre, & Mintz, 2012).

Observational studies (Allen et al., 2013; Westergård, Ertesvåg, & Rafaelsen, 2019) indicate that teachers generally score higher on emotional support and classroom organization than on instructional support. Previous research findings that teachers perceive themselves as skilled in emotional support and classroom organization (Ertesvåg, 2011) are at least partly supported by classroom observations (Allen et al., 2013; Bell et al., 2012; Gitomer et al., 2014; Virtanen et al., 2019; Westergård et al., 2019) and student ratings (Ertesvåg, 2009). Given that teacher–student classroom interaction is considered a multidimensional concept (Hamre et al., 2013; Pianta et al., 2012), instructional support may be influenced by teachers' emotional support and classroom organization (Hamre et al., 2013). High correlations between the three domains have been found in student reports (Ertesvåg & Havik, 2018) and observer scores (e.g., Westergård et al., 2019; Virtanen, et al. 2017; Sandilos, DiPerna & The Family Life Project, 2014). Despite their relatively high correlation, the three domains have been found to be distinct aspects of a three-dimensional concept of teacher–student classroom interaction (e.g., Ertesvåg & Havik, 2018; Hafen et al., 2015). Accordingly, it is reasonable to assume that scores on instructional support may be affected by scores on emotional support and classroom organization as teachers' affective relationship with the individual student may influence student engagement (Ruzek et al., 2016) and subsequent interest and willingness to participate in learning activities. Similarly, teachers' ability to develop effective structures and routines and to prevent misbehavior may affect their ability to provide instructional support (Pianta et al., 2012).

Teachers' trajectories in instructional support

Previous research on lower secondary school teachers' improvement in instructional quality is scarce in general, and it is particularly scarce with regard to instructional support. Generally, studies find that beginning teachers perform worse on measures of instructional quality than experienced teachers (Dicke, Schmeck, Elling, & Leutner, 2015; McCarthy et al., 2015; Schmidt, Klusmann, Lüdtke, Möller, & Kunter, 2016) and that it takes about six years to establish one's teaching style, with growth being largest in the three first years (Brekelmans, Wubbels, & van Tartwijk, 2005). Given that schools contain teachers with varying years of work experience, it can be expected that there are differences in the development of instructional support over time, regardless of interventions. Additionally, improvement may occur as a result of systematic interventions aimed to improve teachers' skills (Allen, Hafen, Gregory, Mikami, & Pianta, 2015; Korpershoek, Harms, de Boer, van Kuijk, & Doolaard, 2016).

Looking at the developmental trajectories of instructional support skills, the possibility of a decrease in performance over time cannot be ruled out. Huberman (1993) described, for example, five consecutive stages of a teacher's career (survival and discovery; stabilization, experimentation/activism and stocktaking; serenity and conservatism; and disengagement), which are closely connected to individual teaching experiences. These stages represent major phases of teachers' development (Richter, Kunter, Klusmann, Lüdtke, & Baumert, 2011). Because teachers are at different stages of their careers, their development may differ; the possibility that there is both positive and negative development among teachers cannot be ruled out.

Implementation theory and previous research (e.g., Ertesvåg, 2009; Ertesvag & Vaaland, 2007; Fullan, 2015; Hall & Hord, 2015) indicate that implementing school-wide initiatives takes longer than one and a half years. Consequently, improvement in instructional support among all teachers was not expected. However, it was assumed that based on, e.g., differences in implementation quality (e.g., Blase, Van Dyke, Fixsen, & Bailey, 2012; Domitrovich et al., 2008) or as part of beginning teachers' normal development as a teacher (Brekelmans et al., 2005), some teachers may improve their instructional support skills. Moreover, it is possible that teachers' job satisfaction, emotional exhaustion and/or stress may affect their performance, regardless of their work experience.

Teacher job satisfaction, emotional exhaustion, stress, and instructional support

Job satisfaction, which is understood as sense of fulfillment and pride felt by people who enjoy their work (Caprara, Barbaranelli, Borgogni, & Steca, 2003; Klassen, 2010; Klassen, Usher, & Bong, 2010; Locke, 1976), is important for teachers because it may influence their performance (Arens & Morin, 2016; Jennings & Greenberg, 2009) and involvement in PD (Gorozidis & Papaioannou, 2014; Nir & Bogler, 2008). High levels of stress and emotional exhaustion among both beginning (Dicke et al., 2015; Hultell, Melin, & Gustavsson, 2013; Klassen & Chiu, 2011) and experienced teachers (Klassen & Chiu, 2011; McCarthy et al., 2015) is of concern because of their potential influence on job satisfaction and performance. A substantial amount of previous research has investigated the negative consequences of teacher stress and emotional exhaustion on work performance (e.g., Halbesleben & Bowler, 2007; Klassen & Chiu, 2011) and the intention to quit among both practicing (Skaalvik & Skaalvik, 2011; Spilt, Koomen, & Thijs, 2011) and pre-service teachers (Klassen & Chiu, 2011). In an early review of previous research, Maslach, Schaufeli, and Leitner (2001) concluded that emotional exhaustion is related to lower productivity and less effectiveness at work. They also hypothesized that emotional exhaustion is not just passively experienced – rather, it prompts individuals to actively distance themselves from their work both emotionally and cognitively.

Spilt et al. (2011) argue that teachers have a basic need for relatedness with the students in their class. Student misbehavior is particularly pertinent for beginning teachers who lack classroom experience and thus do not have well-developed coping strategies in place (Evertson & Weinstein, 2006; Schmidt et al., 2016). However, problems regarding teacher–student relationships, and disruptive student behavior not only are important causes of attrition among beginning teachers (Spilt et al., 2011) but also may cause teachers stress and exhaustion later in their career (Evertson & Weinstein, 2006; Hultell et al., 2013). On the other hand, positive teacher–student relationships may be positively related to teachers' job satisfaction (Grayson & Alvarez, 2008; Sinclair, Dowson, & McInerney, 2006).

There is little research on how improvement in teacher–student relationships is associated with teachers' job satisfaction, stress, and exhaustion (Spilt et al., 2011). The literature suggests that job satisfaction, emotional exhaustion and stress may differ between teachers who improve in instructional support and those who do not. Teachers' job-related emotional exhaustion and stress may be concurrently related to their quality of instruction over time (Chang, 2009; Jennings & Greenberg, 2009). When

teachers have difficulty providing high-quality support, it may be concurrently related to their emotional exhaustion and stress, which make them perceive their workload as more burdensome and unmanageable (Pas, Bradshaw, & Hershfeldt, 2012). Similarly, when teachers are emotionally exhausted, they may struggle to build supportive relationships with their students, sensitively support students' learning and effectively redirect student misbehavior (Pas et al., 2012). Job satisfaction has been found to be higher among teachers performing high in all aspects of teacher–student interaction than among teachers with low performance in this area (Virtanen et al., 2019).

Assessing whether change in instructional support co-occurs with job-related factors (e.g., job satisfaction, emotional exhaustion, stress) can aid in the identification of targets for interventions to support teachers in their efforts to provide high-quality instructional support.

Evaluation of improvement in instructional support

Recently, there has been increasing concern regarding the lack of research uptake in classrooms and schools in general, regardless of the field (e.g., Cooper, Levin, & Campbell, 2009; Harris et al., 2013). Many practitioners are coaxed into accepting improvement strategies, approaches and packages supported by the thinnest veneer of research evidence (e.g., Harris et al., 2013; Simmons, 2011). Additionally, the process from input of research evidence through teachers' learning process with the intent of increasing teacher knowledge, improving classroom practices, improving student learning processes, and, finally, improving student outcomes is complex (Vermunt, 2011, 2014). Even when high-quality research knowledge is systematically and effectively implemented by motivated teachers, the challenges of improved teaching quality are evident.

The impact of PD on the relationship between teacher support and student outcomes is increasingly evaluated (Allen et al., 2015; Gregory, Allen, Mikami, Hafen, & Pianta, 2014). Typically, these impact studies address changes in student behavior (see, e.g., Bradshaw, Waasdorp, & Leaf, 2012; Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011; Korpershoek et al., 2016 for an overview) rather than changes in teacher behavior or learning. Most address individual teachers' improvement, and few are school-wide (Ertesvåg, 2009, 2011; Sørlie & Ogden, 2007). These studies are important as they add to the research on specific standardized PD interventions.

Numerous initiatives for improving teacher–student interaction in the classroom are initiated as part of teacher-, school-, district- or national-level initiatives to improve teacher practices rather than as part of *one* specific intervention. In addition to qualitative evaluations (e.g., Dogan, Pringle, & Mesa, 2016; Vangrieken, Meredith, Packer, & Kyndt, 2017) of certain aspects of different forms of teacher PD, there is a need for a stronger emphasis on monitoring change in teachers' practice throughout these initiatives.

Studies must focus on teachers' learning and improvement and factors that can facilitate teachers' learning and engagement with PD (Opfer & Pedder, 2011; Wayne, Yoon, Zhu, Cronen, & Garet, 2008), as the link between teacher PD and student outcomes is not automatic (Cumming, 2002; King, 2014). A first step may be to investigate change in teachers' practice and its correlates in order to improve the

understanding of teacher learning. It is of interest to investigate teachers' improvement in knowledge and skills as part of an evaluation of a PD initiative.

Typically, classroom interactions understood in line with the TTI framework are evaluated using systematic classroom observation (see Hamre et al., 2013; Pianta, 2016 for overview). External classroom observations are time-consuming and usually fail to reflect teachers' typical practices throughout the entire academic year (Araujo, Carneiro, Cruz-Aguayo, & Schady, 2014). Thus, other approaches, e.g., teacher self-reports, may contribute to a fuller picture. Teachers who can accurately self-assess their strengths and weaknesses as practitioners have sufficiently developed reflective capacity (Vanhoof, Van Petegem, Verhoeven, & Buvens, 2009). Self-reports are an effective approach to understand teachers' beliefs about classroom practice. It is through this classroom practice that teachers build the self-awareness they need to critically reflect on everyday teaching within the classroom or after class. Such self-awareness and abilities in self-inquiry and critical thinking are central to continuous professional growth (Hu, Chen, & Fan, 2018). Accordingly, self-reports may provide additional knowledge important for the purpose of developing conduct effective improvement efforts. It should be noted that despite their strengths, teacher self-reports have limitations. Muijs (2006) argues that teachers' self-reports are unreliable as measures of teacher behavior and that measures of beliefs, values, opinions, etc., are more relevant. Accordingly, teachers' self-reports are not suitable for evaluating teacher performance. Still, teachers‘ self-reports may provide important information of their *perceptions* that students and observers may not. The discussion of whether teachers, student and/or observers are valid and reliable sources of information is extensive, and the results are at least partly contradictory (Desimone, Smith, & Frisvold, 2010; Reddy, Dudek, Fabiano, & Peters, 2015; Wagner et al., 2016). Specifically in professional development initiatives, one may argue that teachers‘ perceptions of classroom processes may be important information for their continuous development.

Context of the study

The Norwegian Ministry of Education established the national, school-wide, initiative to develop lower secondary schools in the fields of and classroom management from. All Norwegian lower secondary schools were divided into five groups, with a new group of schools starting each year. Each school participated for a year and a half and chose one or two of the four topics as their main intervention field. In addition, irrespective of the chosen field, all schools focused on improving aspects of the school as an organization. Each school received support (workshops, seminars, etc.) and supervision from a teacher training institution in their region. Moreover, national centers in the respective fields provided expertise and knowledge to all participating schools through national and regional workshops, lectures, guidance, meetings and professional feedback to key personnel (e.g., school leaders and 1–2 resource teachers that were targeted to support colleagues in the implementation) at each school and their school district. Instructional support as part of the wider concept of classroom interaction was addressed in this support. The schools working on classroom management were guided by a document describing theoretical perspectives on classroom management (Postholm, Midthassel, & Nordahl, 2012). However, the document was expected to be

supplemented by recent research in the field over the five years of the initiative. This national initiative provides a context in which teachers' improvement of instructional support can be explored.

Research aims and hypotheses

The aims of the present study were to (a) explore instructional support trajectories perceived by teachers participating in a national initiative that aimed to develop lower secondary schools, (b) explore the roles of emotional support, classroom organization and work experience in the improvement of instructional support and, (c) describe the differences between subgroups of teachers in terms of instructional support trajectories related to job satisfaction, emotional exhaustion and stress.

Based on the theoretical outline above, it is reasonable to assume that due to the relatively short timeframe of the intervention and/or differences in work experience, there are differences in teachers' instructional support trajectories. Moreover, due to the multidimensional nature of teacher–student interaction in the classroom, these differences may be affected by teachers' emotional support and/or classroom organization. Differences in instructional support trajectories between groups of teachers may be related to both differences in stress and exhaustion (which may affect their ability to participate in PD initiatives) and to a general lack of commitment to their work.

The hypotheses were as follows:

(1) There are meaningful qualitative differences in teachers' trajectories of instructional support that can be attributed to the existence of unobserved subgroups of teachers, when taking emotional support, classroom organization and work experience into consideration.
(2) Emotional support, classroom organization, and work experience are positively related to improvement in instructional support skills.
(3) Teachers who improve in their instructional support experience higher job satisfaction and less emotional exhaustion and stress than teachers who decline or do not improve in instructional support.

The theoretical model is visualized in Figure 1. Additionally, an approach relevant for evaluating improvement in teachers' practice in professional development is discussed.

Methods

Sample and procedure

The sample consisted of 227 teachers at nine participating schools that had chosen classroom management as a focus area. All teachers in these schools were invited to participate in a survey three times: at the beginning of the initiative, which was at the start of the school year (T1), at the end of the same academic year (T2), and after the schools' participation in the initiative ended (T3), which was after one and a half academic years. The response rates at the three waves were 72%, 49% and 52%, respectively.

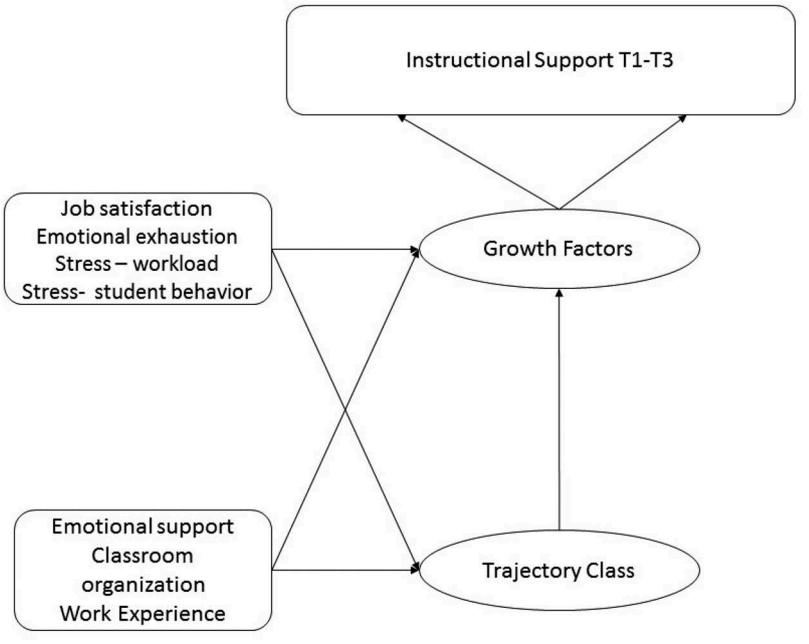

Figure 1. Theoretical model of instructional support trajectories with covariates.

Measurement

Instructional-, emotional- and organizational support

Instructional support was measured by a five-item scale developed for the current study. The scale was a teacher version of a previously developed and validated student measure (Ertesvåg, 2016; Ertesvåg & Havik, 2018). The items addressed how teachers described clear learning targets, helped students understand concepts and facts, asked challenging questions, gave feedback and encouraged discussions that extended students' knowledge. The items were rated on a 6-point scale ranging from 0 ('not at all') to 5 ('completely true'). The mean scores of the items at each time point were included in the analysis. The instrument was originally developed to correspond to the key elements of instructional support described by Pianta et al. (2012) and corresponded to the instructional support domain in the Secondary Version of the Classroom Assessment Scoring System (CLASS-S) (Pianta et al., 2012). Cronbach's alphas were 0.80, 0.78 and 0.77, respectively, for data waves 1–3.

Emotional support and *organizational support* were measured with two four-item scales (Ertesvåg, 2011) in the first wave of data. The two scales measure warmth and control as part of an authoritative teaching style. The concepts are similar to emotional support and classroom organization in TTI. Examples of items on the warmth scale are 'I work actively to create good relationships with my pupils' and 'I show interest in each pupil'. Examples of items in the control scale are 'I have established routines/rules for how the pupils are

supposed to act when they change activity/workplace, etc.' and 'I am closely monitoring the pupils' behavior in class'. The scores for both scales ranged from 0 ('never') to 5 ('very often'). The scales have been previously used and validated (Ertesvåg, 2011). In presenting the results and the discussion, they are referred to as warmth and control, which accurately describe what this scale measures. Warmth corresponds well with the aspects of positive climate and teacher sensitivity in Pianta and colleagues' (2012) conceptualization of emotional support, but they correspond less to regard for adolescent perspectives. Control corresponds well with the domains of behavioral management and negative climate in classroom organization, but they correspond less to productivity. Cronbach's alpha was 0.80 for warmth and 0.78 for control.

Job satisfaction and well-being

Four measures were used to measure teachers' job satisfaction, emotional exhaustion and stress at work at the first wave of data.

Job satisfaction was measured using a slightly modified version of the work satisfaction scale by Starnaman and Miller (1992), which includes five items. The modified version has been used in Norwegian studies both as a five-item version (Westergård, 2007) and as a four-item version (Munthe, 2003). Sample items are 'I experience my work as useful' and 'My work gives me an experience of satisfaction'. Cronbach's alpha was 0.90.

Emotional exhaustion was measured with a five-item scale (Maslach & Jackson, 1981) capturing teachers' emotional well-being at work. Sample items are 'I feel emotionally drained' and 'I feel burned out as a teacher'. Cronbach's alpha was 0.87.

The job satisfaction and emotional exhaustion scales were rated on a scale from 0 ('strongly disagree') to 5 ('strongly agree').

Stress was measured using seven items from the Teacher Stress Inventory (Boyle, Borg, Falzon, & Baglioni, 1995). These items represented two factors: stress from workload and stress from student behavior. An additional item regarding class size was also incorporated, as suggested by Gates (2007). The eight items were previously used by Klassen (2010).

Cronbach's alpha was 0.67 for stress from workload and 0.90 for stress from student behavior.

The respondents answered each item on the two stress scales ranging from 1 ('no stress') to 9 ('extremely stressful').

Teachers' work experience

Teachers reported their years of work experience. Based on previous research indicating that it takes five to six years for teachers to find their teaching style (Brekelmans et al., 2005), teachers were divided into two groups: those with six or fewer years of work experience and those with more than six, coded 1 and 2, respectively.

For all scales used, the mean score of the items included was applied.

Analysis of data

Factor analysis was used to reduce the data from the questionnaires into a number of robust and meaningful underlying factors for instructional support, warmth and

control. Growth mixture models (GMM) were used to identify meaningful population subgroups that were characterized by similar within-group trajectories of change over time (Muthèn & Muthèn, 1998–2017; Wickrama, Lee, O'Neal, & Lorenz, 2016). Establishing how many distinct patterns of change exist in the data is a crucial step in any GMM analysis. It was assumed that teachers' warmth, control, and work experience may affect the number and composition of latent classes (Grimm & Ram, 2009; Wickrama et al., 2016). Accordingly, the direct specification approach (i.e., a 1-step approach) (Wickrama et al., 2016) to add covariates was applied to allow these covariates to affect teacher class membership (Figure 1). The results indicated that warmth did not have an impact on class membership when control already was included, and thus warmth was omitted from further analysis.

The choice regarding the extraction of latent classes was guided by information criteria such as the Bayesian information criterion (BIC), the Akaike information criterion (AIC), and the sample-sized adjusted BIC (SSABIC), likelihood-ratio tests such as the adjusted Lo-Mendell-Rubin likelihood-ratio test (A-LRT), and the parametric bootstrap method (BLRT) test. Information criteria were scaled such that smaller values indicated a better fit, whereas the A-LRT and BLRT tests compared the improvement in fit between a model and a model with one less class (e.g., a two-class model versus a one-class model) to identify whether adding a class significantly improved the model (Wickrama et al., 2016).

Conventional analysis was conducted using SPSS 25.0, and GMM analysis was conducted using Mplus version 8.0. Little's (1988) MCAR test indicated that missingness was completely at random: $\chi^2(313) = 296.60$, $p = .739$. Full information maximum likelihood with robust standard errors was used to estimate the model's parameters as it permits the inclusion of all available data in the analysis. This strategy for handling missing data is considered an appropriate method of modeling with missing data (Berlin, Parra, & Williams, 2014; Little, Jorgensen, Lang, & Moore, 2014). Missing information for covariates was generally at low levels (<5% for most items). Nonetheless, the cumulative effects of missing information meant that a complete case analysis (N = 187) contained 15% fewer cases than the full sample. An inspection of the cases that were excluded because of missingness did not reveal any significant differences in the key variables included in the study.

Results

The aim of the study was to explore teachers' improvement in instructional support throughout an intervention that addressed classroom management, to examine the influence of control and work experience and to identify differences in job satisfaction, emotional exhaustion and teacher stress.

Identifying the number of latent classes

Although there was reason to expect more than one trajectory class based on theory and previous studies, a specific number of trajectory classes was not expected. There may be groups of teachers perceiving change in instructional support, for example, based on

implementation quality or work experience. These change patterns are not necessarily similar. Therefore, as part of the data exploration, one to four classes were tested.

Descriptive statistics for variables included in the model are shown in Table 1. The model fit for GMM of instructional support is shown in Table 2. As indicated by the values in bold for the best fit for each information criterion, the indices were somewhat contradictory with respect to the number of classes. In general, the BIC and SSABIC statistics have been reported as 'better' performance indices than the AIC statistics (Peugh & Fan, 2012; Wickrama et al., 2016). The BIC indicates that a one-class model fits the data best, with the fit becoming increasingly worse as the number of classes increases. The SAABIC indicates that a four-class solution fits the data best. The AIC model fit statistics continued to improve as the number of classes increased from 1 to 4. Entropy was highest for a four-class model, approaching the value of 0.80 considered high entropy (Wickrama et al., 2016). The LMR-LRT and the adjusted LMR-LRT indicated that a two-class model did not fit significantly better than a one-class model. The BLRT indicated a two-class model. Moreover, for the two-trajectory model, the probability of correct classification into group membership was 0.90 for one of the trajectory classes and 0.95 for the other trajectory class, indicating that the

Table 1. Descriptive statistic for variables included in the model. Mean, standard deviation (SD), minimum (Min) value, maximum (Max) value, skewness (Skew), and kurtosis (Kurt).

	Mean	SD	Min	Max	Skew	Kurt
Instructional support T1[a] (n = 187)	3.83	0.59	1.80	5.00	−0.227	0.098
Instructional support T2 [a] (n = 111)	3.93	0.51	2.80	5.00	0.103	−0.629
Instructional support T3 [a] (n = 116)	3.93	0.52	2.80	5.00	0.276	−0.482
Control[b] (n = 187)	4.20	0.54	2.67	5.00	−0.213	−0.528
Work experience[c] (n = 187)	1.76	0.47	1.00	2.00	−1.213	−0.528
Job satisfaction (n = 187)	4.29	0.69	0.00	5.00	−1.682	7.104
Emotional exhaustion	3.36	1.12	0.00	5.00	−0.658	0.078
Stress due to workload	5.10	1.73	1.00	9.00	−0.308	−0.504
Stress due to student behavior	3.86	1.87	1.00	8.75	0.256	−0.926

a. Score range from 1 to 5, high score is positive. b. Score range from 0 to 3, high score is positive. c. Score on work experience indicates the composition of the groups with respect to having up to six years (score 1) or more (score 2).

Table 2. Fit statistics across the classes/subgroups for the GMM of instructional support.

Number of classes	1	2	3	4
	Class1: N = 187	Class1: N = 125 Class2: N= 62	Class1: N = 39 Class2: N = 34 Class3: N = 114	Class1: N = 31 Class2: N = 32 Class3: N = 83 Class4: N = 43
AIC	495.86	480.18	478.16	**473.19**
BIC	**524.49**	541.57	568.63	592.74
Sample-size adjusted BIC	492.82	481.39	479.94	475.55
Entrophy	NA	**0.76**	0.73	**0.77**
VUONG-LO-MENDELL-RUBIN LIKELIHOOD RATIO TEST FOR 1 (H0) VERSUS 2 CLASSES p-value	NA	0.086 0.092	0.308 0.317	0.053 0.055
LO-MENDELL-RUBIN ADJUSTED LRT TEST p-value				
BLRT p-value	NA	**0.000**	0.428	0.333

Note. GMM = growth mixture modeling; AIC = Akaike information criterion; BIC = Bayesian information criterion. BLRT = bootstrapped likelihood ratio test.

two identified trajectories successfully grouped teachers with similar within-trajectory patterns of change over time. Although some fit indices indicated a four-class model and the BIC indicated a one-class model, a two-class model was selected. There were both statistical and theoretical reasons for this decision. The preference for the more parsimonious model, the model with fewer classes, when fit indices are similar (Feldman, Masyn, & Conger, 2009) guided the choice for a two-class model over a four-class model. This was supported by the plot indicating that the four-class model contained three classes with stable trajectories but different levels of instructional support, and it contained a fourth improving class. Accordingly, it provided a similar pattern as the two-class model except that there were more non-improving trajectories. Additionally, the choice of the two-class model was supported by the posterior probability scores, which indicated a high probability of correct classification. The preference for a more parsimonious model supports the argument for a one-class model over a two-class model. However, as argued above, there are theoretical and empirical reasons to expect that some teachers change over time, even if the majority do not change significantly in a relatively short time period. The classes should be clearly interpretable and consistent with theory and should not be based solely on model fit indices (Wickrama et al., 2016). Accordingly, based on a combination of model fit indices, theory and previous research, the two-trajectory model was selected as the final model.

Description of the two latent classes

The two-trajectory model indicated two types of change patterns. The largest class (containing 127 teachers) displayed a non-significant decrease in instructional support skills ($\beta = -0.014$, $p = 0.640$) (Hereafter, this group is identified as the normative group). A smaller group (62 teachers) indicated significant improvement ($\beta = 0.19$, $p = 0.000$) over time (hereafter, the improving group). The improving group had a higher intercept ($M = 3.99$) than the normative group ($M = 3.75$); this indicates that at the start of the initiative, the teachers in this group already had higher skills than teachers in the normative group. Initial control skills and work experience both contributed to predicted trajectory patterns. Teachers in the improving group reported a higher level of control ($M = 4.67$) than those in the normative group ($M = 3.97$). Similarly, the results indicated that the normative group contained a slightly higher number of experienced teachers ($M = 1.79$) than the improving group ($M = 1.70$). The trajectories for the two groups are shown in Figure 2.

Characterization of within-class trajectories of change

Differential patterns of predicted variations in the latent growth factors of the two latent classes provide a more accurate depiction of these classes (see Table 3). For the normative group, teacher control significantly predicted the initial score on instructional support. A 1-point increase in the teacher control score was associated with, on average, a 0.68 ($p = 0.000$) higher score in instructional support, all else being constant. The results indicate that in the normative group, teachers who reported high control also tended to report a high initial level of instructional support. Moreover, compared

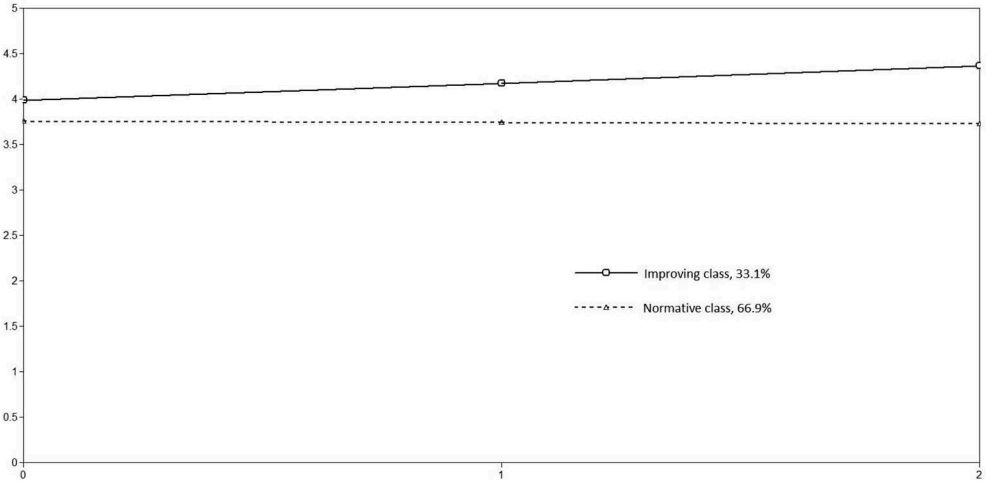

Figure 2. Estimated means instructional support trajectories for the improving class and the normative class.

Table 3. Characterization of within-class trajectories of change in instructional support.

			β	p.
Normative class	I	Control	0.68	.000
		Work experience	−0.18	.060
	S	Control	−1.03	.000
		Work experience	0.16	.362
Improving class	I	Control	0.81	.000
		Work experience	−0.01	.933
	S	Control	−1.14	.000
		Work experience	0.45	.020

to teachers with more than six years experience, teachers with less than six years work experiences are expected to score 0.18 higher on average in instructional support, at the same level of control. The association was marginally non-significant (p = 0.060).

The results for the improving group indicate a significant association with teacher control at the outset of the initiative. A 1-point increase in the teacher control score was associated with, on average, a 0.81 higher instructional support score (p = 0.000). There was no significant association with work experience for the improving group (β = –0.014, p = 0.933); this indicated that experienced and less experienced teachers in the improving group did not differ with respect to the initial instructional support score.

For the normative group, the regression of the latent slope on teacher control and work experience indicated that only teacher control significantly predicted the rate of change in instructional support scores. A 1-point increase in the teacher control score was associated with an average decline of 1.03 (p = 0.000) in the rate of change in instructional support scores. The results indicated that the higher the teacher control score, the less steep the change in instructional support scores. This implies that the teachers who scored lowest on control improved the most throughout the initiative. Work experience was not related to change in instructional support for the normative

group (β = 0.16, p = 0.362). For teachers in the improving group, a 1-point increase in the teacher control score was associated with an average decline of 1.13 (p = 0 .000) in the rate of change in instructional support scores. This indicated that the higher the teacher control score, the less steep the score change. The results indicated that work experience was related to change in instructional support for the improving group (β = 0.45, p = 0.020); this indicated that compared to inexperienced teachers, the most experienced teachers improved more on average, at the same level of control.

Job satisfaction, emotional exhaustion and teacher stress

To explore differences in job satisfaction, emotional exhaustion and stress between teachers in the two trajectory groups, the identified class membership was saved and used for further analysis. Table 4 provides descriptive statistics for the two groups regarding emotional exhaustion, job satisfaction, stress due to student behavior and stress due to workload. The results indicate that teachers in the increasing group reported significantly (F(1) = 9.75, p = 0.002) higher job satisfaction than teachers in the normative group. The effect size was medium. Similarly, the improving group reported less stress from student behavior than the normative group. The difference was non-significant (F(1) = 3.53, p = 0.062), and the effect size was small. No significant differences were found between the two groups with respect to emotional exhaustion and stress from workload.

Discussion

The aim of the study was to explore teachers' improvement in instructional support skills throughout a national initiative that aimed to develop lower secondary schools. The study also aimed to investigate the impact of control and work experience as well as to describe differences in job satisfaction, emotional exhaustion and stress between teachers. The GMM approach to evaluating improvement in teachers' practice in professional development is briefly discussed.

Techers' improvement in instructional support

In line with the first hypothesis, the analysis of the heterogeneity of instructional support trajectories indicated that the two subgroups of teachers (i.e., normative and improving) exhibited substantively different levels of change. The majority of teachers

Table 4. Means, standard deviations, and effect sizes (Cohens d) for teacher job satisfaction, emotional exhaustion, stress due to workload and stress due to student behavior.

		Trajectory class		
	Whole sample	Normative class	Improving class	*d*
	M (SD)	M (SD)	M (SD)	
Job satisfaction[ac]	4.29 (0.69)	4.18 (0.68)	4.52 (0.67)	0.50
Emotional exhaustion[bc]	3.36 (1.13)	3.29 (1.04)	3.50 (1.27)	0.18
Stress from workload[ad]	5.10 (1.73)	5.15 (1.64)	4.98 (1.91)	−0.10
Stress from student behavior[ad]	3.86 (1.87)	4.05 (1.86)	3.50 (1.87)	−0.30

a. High score implies high job satisfaction or stress, b. Low score implies high emotional exhaustion, c. Score 0–5, d. Score 1–9

did not report a substantial change in instructional support throughout the initiative. However, one group of teachers reported that they had improved in their instructional support skills. The lack of change in instructional support for most teachers may be explained by the relatively short timeframe of the initiative. Based on implementation research (e.g., Blase et al., 2012; Greenberg, Domitrovich, Graczyk, & Zins, 2005; Hall & Hord, 2015), it is not expected that all teachers will improve within one and a half years. Moreover, as Bubb, Earley, and Hempel-Jorgensen (2008) found, not all teachers feel obligated to improve from PD. This may be particularly true for whole-school initiatives, where individual teachers are not directly involved in the decision to take on the initiative and may be reluctant, hesitant or critical (Rogers, 2002, 2003). Similar to the current initiative, in the international educational arena, improvement efforts in school are often introduced as centrally organized in-service training programs or continuous PD programs. In many cases, participation in these programs is optional, and when it is mandatory, there is no way of ensuring teachers' optimal engagement in the initiatives. Teachers' autonomous motivation predicts their intention to participate in training and to devote effort to improving, whereas controlled motivation does not (Gorozidis & Papaioannou, 2014). Teachers' willingness to learn (i.e., motivation) is one of the basic features of teacher learning and successful PD (Shulman & Shulman, 2004). Accordingly, motivation to learn may be an important factor determining teachers' improvement. This may be particularly true for instructional support, which is a key element of classroom management. Over at least two decades, Norwegian policy documents have strongly emphasized teachers' instructional role and classroom management (e.g., Ministry of Education, 2011, 2016). Teachers have been involved in a substantial number of PD efforts in related areas, although they do not specifically address instructional support. Previous research indicates that teachers perceive themselves as skilled in related areas such as authoritative teaching (Ertesvåg, 2011). These findings are supported by student reports (e.g., Bru, Stornes, Munthe, & Thuen, 2010; Ertesvåg, 2009) and classroom observations (Westergård et al., 2019).

It should be noted that although high scores have been found in studies using diverse respondents, the variation was larger in student reports and observations than in self-reports. Of particular interest is that instructional support measured by the Classroom Assessment Scoring System (CLASS; Pianta et al., 2012) indicated lower scores for aspects of instructional support than for emotional support and classroom organization (Westergård et al., 2019). By contrast, teachers in the current study reported relatively higher scores. Although there was some variation in the findings between the respondent groups, teachers, in general, are found to be skilled in key aspects of classroom management. This may be a reason for the lack of improvement among most teachers: they already felt skilled in this area and therefore may have been less motivated to invest time and resources in PD in this area. The findings illustrate an advantage of teachers' self-reports in the evaluation of PD. Although observations and student reports may suggest lower teacher instructional support, the teachers' own perception may provide contradictory information. The perception of their own performance may provide important knowledge before a change effort is implemented. It is not surprising if teachers are not motivated to work on development in a field they experience as mastered.

It has been suggested that there are thresholds for teacher–student interactions and student outcomes (Pianta, 2016). This indicates that at a certain level of consistency

and/or stability, students' exposure to high-quality teacher–student interactions benefits student outcomes. If teachers, for whatever reason, do not perform at high level with a certain consistency, level and stability, it may affect improvement in instructional support. Another possible reason for the findings may be, as noted above, that even if teachers are motivated, one and half years is not enough time to fully implement the new knowledge in the classroom in a way that reflects teacher mastery of new skills (see Hall & Hord, 2015). Although the majority of the teachers did not report improvement, one group did. A possible reason for this may be differences in implementation quality among teachers and/or schools (Blase et al., 2012; Greenberg et al., 2005; Hall & Hord, 2015). Another possible reason may be teachers' natural development (Brekelmans et al., 2005), which, accordingly, is not a result of the initiative. However, the finding that the most experienced teachers improved the most does not support this notion.

The impact of control and work experience

The second hypothesis was partly confirmed, as control was positively related to initial status in instructional support skills for both the improving group and normative group, but was negatively associated with rate of change in instructional support in both groups. Work experience was significantly associated with change in the improving group only.

Previous research has indicated that it takes about six years for a teacher to establish his/her teaching style, and teachers develop the most in the first three years (e.g., Brekelmans et al., 2005). However, given that the most experienced teachers reported the highest improvement, this was not the explanation for the improving group. Work experience did not predict teachers' initial status in this group. Accordingly, the increase was not explained by lower initial scores for experienced teachers than for inexperienced teachers. The result controls for the effect of control skills among teachers. A previous study found a high correlation between instructional support and classroom organization, which is a phenomenon similar to control (Ertesvåg & Havik, 2018). It is possible that teachers performing high in control feel more confident and certain (Munthe, 2003) and consequently improve more in instructional support. This result indicates that among teachers with the same level of control, more experienced teachers improve more. These results are similar to the results indicating that the improving group had a higher initial score. Teachers who scored higher on initial instructional support improved the most, and the teachers with the best performance became even better. It should be noted that on average, teachers in both the normative and improving groups scored relatively high at the beginning of the initiative. However, as the results indicate, control predicted initial status in both groups, and work experience predicted initial status only in the improving group.

Job satisfaction, emotional exhaustion and stress

Given the association between challenges in classroom management and teachers' job satisfaction and stress (McCarthy et al., 2015 for overview), it is not unexpected that teachers in the improving group had higher job satisfaction and less stress than teachers in the normative group. Teachers' job satisfaction and stress from student behavior may

influence their willingness and ability to take on the work needed to improve their teaching skills and may also influence their involvement in PD. However, the research design of the current study does not allow for conclusions about causality. It is not known whether teachers in the improving group improved more than teachers in the normative group because they perceived higher satisfaction and less stress in their job or they were more satisfied and less stressed because they perceived that their teaching was high quality. A previous study found that teachers performing high in all aspects of classroom interaction, as scored by external observers, reported higher job satisfaction than teachers scoring lower on all aspects (Virtanen et al., 2019); this indicates that there is a relationship between profiles of classroom interaction, including instructional support and job satisfaction. More research is needed to address teachers' instructional support skills and their ability to maintain job satisfaction and health by preventing stress and burnout.

Moreover, studies of the interaction between teachers' self-efficacy (in classroom management) and stressors (e.g., student misbehavior) (see Dicke et al., 2015; Pianta, 2016 for an overview) may elucidate the results of the current study. Self-efficacy is understood as a person's belief in, or perception of his/her capability to successfully undertake the actions required to complete a given task (Bandura, 1997). Teachers with high self-efficacy expect that they are able to improve students' behavior and achievement despite difficulties such as adverse environmental influences. In contrast, teachers with low self-efficacy expect that they have little capacity to have an impact on students' motivation and academic and social outcomes (see Bandura, 1997; Pianta, 2016; Tschannen-Moran & Woolfolk Hoy, 2001). Moreover, teacher self-efficacy has been positively linked to teacher commitment (Klassen & Chiu, 2011), job satisfaction (Caprara et al., 2003; Dicke et al., 2015) and instructional quality (Tschannen-Moran & Woolfolk Hoy, 2001). Accordingly, teachers' perceived ability to impact instructional support may be associated with their job satisfaction and stress from student behavior.

Methodological consideration

The strengths of the study lie in the longitudinal design and the statistical methodology used (GMM). Clearly, the study has some limitations that should be addressed. First, the lack of a comparison group did not allow for an investigation of the effects of the initiative. Nonetheless, it was possible to investigate improvement resulting from natural development and/or training and to examine the covariates thereof. Moreover, the sample was not random. However, teachers at both urban and rural as well as small and large schools were included. Nonetheless, the study would benefit from a comparison group and a random sample, which should be included in future studies.

Teacher self-reports as the only source of information is a limitation of the current study. Future studies of implementation of PD in instructional support may benefit from eliciting a variety of respondents, including teachers, students, and external observers, to provide complementary knowledge.

That said, the approach illustrates how PD initiatives may be monitored by identifying perceived trajectories of change. The main aim is not to generalize the results and add to the knowledge on improvement in general but to monitor change in an improvement initiative, in a specific intervention or an ongoing PD, which are usually

not evaluated. As such, the approach may supplement, e.g., small-scale qualitative investigations. An improved understanding of the measurements to identify improvement in both specific initiatives and ongoing PD may improve the ability to monitor teachers' uptake of research in both practice and research.

Although not a limitation, the approach to GMM should be noted. A 1-step approach (Wickrama et al., 2016) to add covariates was chosen over a 3-step approach. In the research literature, there is some debate regarding the timing of the inclusion of covariates (see, e.g., Li & Harring, 2017 for overview). Previous research has suggested that covariates of latent group membership may be included when deciding on the number of latent classes (e.g., Li & Harring, 2017; Lubke & Muthén, 2007), while others have advocated their inclusion only after class enumeration (e.g., Nylund-Gibson & Masyn, 2016). The 1-step approach was chosen based on the theoretical assumptions and is supported by previous research. Instructional support is part of the multidimensional concept of classroom interaction (see outline of the TTI framework above). Accordingly, it was expected that emotional support and classroom organization may affect class membership, and they were included in the analysis from the beginning. Future research may shed further light on the debate on the use of different approaches to GMM.

A note on missingness must be added. Although it did not meet the textbook suggestions, the response rate was acceptable considering the response rate of earlier studies of teachers (e.g., Ertesvåg, 2011; Midthassel, Bru, & Idsøe, 2000; Munthe, 2001). At the sample level, the response rate, especially in the second and third waves, was low. The possibility that the samples were somewhat biased cannot be ruled out. It is not uncommon for there to be lower-than-desirable response rates in longitudinal designs (e.g., Idsoe, Hagtvedt, Bru, Midthassel, & Knardahl, 2008; Mäkikangas, Feldt, & Kinnune, 2006). Given that the study was conducted over one and a half years, some of the attrition was due to teacher turnover. Previous results have indicated that the least experienced teachers are more likely to drop out, most likely because beginning teachers are less likely to have a permanent position (Ertesvåg, 2011). This type of attrition did not appear to be caused by the initiative. The application of the maximum likelihood method in the present study allowed for the use of all observations in the data set when estimating the trajectory models. A major advantage of including all data is improvement in the statistical power. The possibility of drawing on all available data was a strength of the study. Additionally, Little's test indicated that missingness was completely random.

Practical implications of the study and conclusions

The current study also offers several significant contributions and implications for educational policymakers, school district leaders, principals and teachers. This study provides support for the previous research indicating that school-wide change takes longer than one and a half years, even in areas teachers generally feel that they have mastered before an initiative. Although the majority of teachers did not report improvement, one third did. A thought-provoking finding is that the teachers who improved reported higher pre-initiative performance than the teachers who did not improve. The teachers performing the best became even better. This may indicate that the initiative did not last long enough to reflect improvement among the lower-performing teachers.

However, it may also be the case that the initiative, with its 'one-size' approach, was not sufficient to facilitate improvement among lower-performing teachers.

In addition to adopting collective or group approaches to PD, it may be beneficial for principals, policymakers and school district leaders to create PD that is tailored to the needs of individual teachers when searching for ways to increase teacher motivation and well-being. PD should provide opportunities for teachers to work together to solve problems, provide and obtain feedback, gain knowledge, and receive individualized supervision. This may, for example, impact the extent to which teachers experiment with new materials and approaches in their classroom, try out new methods, and reflect on their teaching practice, thereby providing teachers with perception of success, encouraging personal investment in work and increasing job satisfaction and well-being.

Most teachers did not improve, and this group of teachers reported the lowest initial instructional support. Given the association between quality of instructional support and student outcomes found in previous research, the overall findings of the current study support the notion that teacher learning is important for improved student learning. Accordingly, teachers' improvement must be monitored in any PD initiative. Teachers' sense of mastery and PD is important not only for their own job satisfaction and well-being but also for their students' engagement and learning outcomes.

Disclosure statement

No potential conflict of interest was reported by the author.

Funding

This work was supported by grants from The Research Council of Norway [Project number 238003/H20].

Notes on contributor

Sigrun K. Ertesvåg is Professor of Educational Psychology and Head of Research at the Norwegian Center for Learning Environment and Behavioral Research in Education, University of Stavanger, Norway. Her research interest are in the area of school improvement, school organization and implementation of professional development for teachers. She has a special interest in teachers' implementation of research based knowledge into practice in their classroom.

ORCID

Sigrun K. Ertesvåg http://orcid.org/0000-0001-7804-8571

References

Allen, J., Gregory, A., Mikami, A., Lun, J., Hamre, B., & Pianta, R. (2013). Observations of effective teacher-student interactions in secondary school classrooms: Predicting student

achievement with the classroom assessment scoring system-secondary. *School Psychology Review*, *42*(1), 76–98.
Allen, J. P., Hafen, C. A., Gregory, A. C., Mikami, A. Y., & Pianta, R. (2015). Enhancing secondary school instruction and student achievement: Replication and extension of the my teaching partner-secondary intervention. *Journal of Research on Educational Effectiveness*, *8* (4), 475–489.
Araujo, M. C., Carneiro, P., Cruz-Aguayo, Y., & Schady, N. (2014). *A helping hand? Teacher quality and learning outcomes in kindergarten*. Washington, DC: Banco Interamericano de Desarrollo. Inedito.
Arens, A. K., & Morin, A. J. S. (2016). Relations between teachers' emotional exhaustion and students' educational outcomes. *Journal of Educational Psychology*, *108*(6), 800–813.
Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: Freeman.
Bell, C. A., Gitomer, D. H., McCaffrey, D. F., Hamre, B. K., Pianta, R. C., & Qi, Y. (2012). An argument approach to observation protocol validity. *Educational Assessment*, *17*(2–3), 62–87.
Berlin, K. S., Parra, G. R., & Williams, N. A. (2014). An introduction to latent variable mixture modeling (Part 2): Longitudinal latent class growth analysis and growth mixture models. *Journal of Pediatric Psychology*, *39*(2), 188–203.
Blase, K. A., Van Dyke, M., Fixsen, D. L., & Bailey, F. W. (2012). Implementation science: Key concepts, themes, and evidence for practitioners in educational psycholo. In B. Kelly & D. N. Perkins (Eds.), *Handbook of implementation science for psychology in education* (pp. 13–36). Cambridge: Cambridge University Press.
Boyle, G. J., Borg, M. G., Falzon, J. M., & Baglioni, A. J., Jr. (1995). A structural model of the dimensions of teacher stress. *British Journal of Educational Psychology*, *65*(1), 49–67.
Bradshaw, C. P., Waasdorp, T. E., & Leaf, P. J. (2012). Effects of school-wide positive behavioral interventions and supports on child behavior problems. *Pediatrics*, *130*(5), 1136–1145.
Brekelmans, W., Wubbels, T., & van Tartwijk, J. (2005). Teacherestudent relationships across the teacher career. *International Journal of Educational Research*, *43*(1–2), 55–71.
Bru, E., Stornes, T., Munthe, E., & Thuen, E. (2010). Students' perceptions of teacher support across the transition from primary to secondary school. *Scandinavian Journal of Educational Research*, *54*(6), 519–533.
Bubb, S., Earley, P., & Hempel-Jorgensen, A. (2008). *Staff development outcomes study*. London: Institute of Education.
Caprara, G. V., Barbaranelli, C., Borgogni, L., & Steca, P. (2003). Efficacy beliefs as determinants of teachers' job satisfaction. *Journal of Educational Psychology*, *95*, 821e832.
Chang, M. (2009). An appraisal perspective of teacher burnout: Examining the emotional work of teachers. *Educational Psychology Review.*, *21*(3), 193–218.
Cooper, A., Levin, B., & Campbell, C. (2009). The growing (But Still Limited) importance of evidence in education policy and practice. *Journal of Educational Change*, *10*(2–3), 159–171.
Cumming, J. (2002). Working together as a profession. *Unicorn*, *28*(2), 1–4.
Desimone, L., Smith, T., & Frisvold, D. (2010). Survey measures of classroom instruction: Comparing student and teacher reports. *Educational Policy*, *24*, 267–329.
Dicke, T., Schmeck, A., Elling, J., & Leutner, D. (2015). Reducing reality shock: Longitudinal effects of a classroom management skill training on pre-service teachers' well being. *Teaching and Teacher Education: an International Journal of Research and Studies*, *48*(May), 1–12.
Dogan, S., Pringle, R., & Mesa, J. (2016). The impacts of professional learning communities on science teachers' knowledge, practice and student learning: A review. *Professional Development in Education*, *42*(4), 569–588.
Domitrovich, C. E., Bradshaw, C. P., Poduska, J. M., Hoagwood, K., Buckley, J. A., Olin, S., … Ialongo, N. S. (2008). Maximizing the implemenaion quality of evidence-based preventive interventions in schools: A conceptual framework. *Advances in School Mental Health Promotion*, *1*(3), 6–28.
Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A metaanalysis of school-based universal interventions. *Child Development*, *82*(1), 405–432.

Ertesvåg, S. K. (2009). Classroom leadership: The effect of a school development programme. *Educational Psychology*, *29*(5), 515–539.
Ertesvåg, S. K. (2011). Measuring authoritative teaching. *Teaching and Teacher Education: an International Journal of Research and Studies*, *27*(1), 51–61.
Ertesvåg, S. K. (2016). Students who bully and their perceptions of teacher support and monitoring. *British Educational Research Journal*, *42*(5), 826–850.
Ertesvåg, S. K., & Havik, T. (2018) *Student mental health problems, proactive aggression and perceived classroom interaction*. Submitted.
Ertesvag, S. K., & Vaaland, G. S. (2007). Prevention and reduction of behavioural problems in school: An evaluation of the respect-program. *Educational Psychology*, *27*(6), 713–736.
Evertson, C. M., & Weinstein, C. S. (2006). *Classroom management as a field of inquiry handbook of classroom management: Research, practice, and contemporary issues* (pp. 3–15). Mahwah, NJ US: Lawrence Erlbaum Associates Publishers.
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.
Feldman, B., Masyn, K. E., & Conger, R. (2009). New approaches to studying problem behaciors: A comparison of methods for modeling longitudinal, categorical adolescent drinking data. *Development Psychology.*, *45*(3), 652–676.
Fullan, M. (2015). *The new meaning of educational change* (5th ed.). NewYork, Routledge: Teacher College Press.
Gates, G. S. (2007). *Emerging thought and research on student, teacher, and administrative stress and coping*. Greenwich, CT: Information Age.
Gitomer, D. H., Bell, C. A., Qi, Y., McCaffrey, D. F., Hamre, B. K., & Pianta, R. C. (2014). The instructional challenge in improving teaching quality: Lessons from a classroom observation protocol. *Teachers College Record*, *116*(6), 1–32.
Gorozidis, G., & Papaioannou, A. G. (2014). Teachers' motivation to participate in training and to implement innovations. *Teaching and Teacher Education: an International Journal of Research and Studies*, *39*(April), 1–11.
Grayson, J. L., & Alvarez, H. K. (2008). School climate factors relating to teacher burnout: A mediator model. *Teaching and Teacher Education: an International Journal of Research and Studies*, *24*(5), 1349–1363.
Greenberg, M. T., Domitrovich, C. E., Graczyk, P. A., & Zins, J. E. (2005). *The study of implemetnation in school-based preventive interventions: Theory, research, and practice*. U.S: Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Centre for Menatal Health Services.
Gregory, A., Allen, J. P., Mikami, A. Y., Hafen, C. A., & Pianta, R. C. (2014). Effects of a professional development program on behavioral engagement of students in middle and high school. *Psychology in the Schools*, *51*(2), 143–163.
Grimm, K. J., & Ram, N. (2009). Nonlinear growth models in Mplus and SAS. *Structural Equation Modeling*, *16*(4), 676–701.
Hafen, C. A., Hamre, B. K., Allen, J. P., Bell, C. A., Gitomer, D. H., & Pianta, R. C. (2015). Teaching through interactions in secondary school classrooms: Revisiting the factor structure and practical application of the classroom assessment scoring system-secondary. *Journal of Early Adolescence*, *35*(5–6), 651–680.
Halbesleben, J. R., & Bowler, W. M. (2007). Emotional exhaustion and job performance: The mediating role of motivation. *Journal of Applied Psychology*, *92*(1), 93–106.
Hall, G. E., & Hord, S. M. (2015). *Implementing change. Patterns, principles and potholes* (4th ed.). LV: Pearson Education Inc.
Hamre, B. K., Pianta, R. C., Downer, J. T., Decoster, J., Jones, S., Brown, J., … Kaefer, T. (2013). Teaching through interactions: Testing a developmental framework of effective teaching in over 4,000 classrooms. *Elementary School Journal*, *113*(4), 461–487.

Hargreaves, A. (2005). Educational change takes ages: Life, career and generational factors in teachers' emotional responses to educational change. *Teaching and Teacher Education*, *21*(8), 967–983.

Harris, A., Chapman, C., Muijs, D., Reynolds, D., Campbell, C., Creemers, B., … Weinstein, J. (2013). Getting lost in translation? An analysis of the international engagement of practitioners and policy makers with the educational effectiveness research base. *School Leadership & Management.*, *33*(1), 3–19.

Hu, B. Y., Chen, L., & Fan, X. (2018). Profiles of teacher-child interaction quality in preschool classrooms and teachers' professional competence features. *Educational Psychology*, *38*(3), 264–285.

Huberman, M. A. (1993). *The lives of teachers*. New York: Teachers College Press.

Hultell, D., Melin, B., & Gustavsson, J. P. (2013). Getting personal with teacher burnout: A longitudinal study on the development of burnout using a person-based approach. *Teaching and Teacher Education*, *32*(May), 75–86.

Idsoe, T., Hagtvet, K. A., Bru, E., Midthassel, U. V., & Knardahl, S. (2008). Antecedents and outcomes of intervention program participation and task priority change among school psychology counselors: A latent variable growth framework. *Journal of School Psychology*, *46* (1), 23.52.

Jennings, P. A., & Greenberg, M. T. (2009). The prosocial classroom: Teacher social andemotional competence in relation to student and classroom outcomes. *Review of Educational Research*, *79*(1), 491–525.

Kane, T., & Staiger, D. (2012). *Gathering feedback for teaching. Combining high-quality observations with student surveys and achievement gains*. MET Project Research Paper, Bill & Melinda Gates Foundation, Seattle, WA. Retrieved from https://eric.ed.gov/?id=ED540960

Kersting, N. B., Givvin, K. B., Thompson, B. J., Santagata, R., & Stigler, J. W. (2012). Measuring usable knowledge: Teachers' analyses of mathematics classroom videos predict teaching quality and student learning. *American Educational Research Journal*, *49*(3), 568.

King, F. (2014). Evaluating the impact of teacher professional development: An evidencebased framework. *Professional Development in Education*, *40*(1), 89–111.

Klassen, R. M. (2010). Teacher stress: The mediating role of collective efficacy beliefs. *The Journal of Educational Research*, *103*(5), 342–350.

Klassen, R. M., & Chiu, M. M. (2011). The occupational commitment and intention to quit of practicing and pre-service teachers: Influence of self-efficacy, job stress, and teaching context. *Contemporary Educational Psychology*, *36*(2), 114–129.

Klassen, R. M., Usher, E. L., & Bong, M. (2010). Teachers' collective efficacy, job satisfaction, and job stress in cross-cultural context. *Journal of Experimental Education*, *78*(4), 464–486.

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). New York: Waxmann Publishing Co.

Korpershoek, H., Harms, T., de Boer, H., van Kuijk, M., & Doolaard, S. (2016). A meta-analysis of the effects of classroom management strategies and classroom management programs on students' academic, behavioral, emotional, and motivational outcomes. *Review of Educational Research.86*, (3), 643–680.

Kunter, M., & Baumert, J. (2006). Who is the expert? Construct and criteria validity of student and teacher ratings of instruction. *Learning Environments Research*, *9*(3), 231–251.

Li, M., & Harring, J. R. (2017). Investigating approaches to estimating covariate effects in growth mixture modeling: A simulation study. *Educational and Psychological Measurement*, *77*(5), 766–791.

Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, *83*(404), 1198–1202.

Little, T. D., Jorgensen, T. D., Lang, K. M., & Moore, W. G. (2014). On the joys of missing data. *Journal of Pediatric Psychology*, *39*(2), 151–162.

Locke, E. A. (1976). The nature and causes of job satisfaction. In M. Dunette (Ed.), *Handbook of industrial and organizational psychology* (pp. 1297–1349). Chicago: Rand-McNally.
Lubke, G. H., & Muthén, B. O. (2007). Performance of factor mixture models as a function of model size, covariate effects, and class-specific parameters. *Structural Equation Modeling*, *14*, 26–47.
Mäkikangas, A., Feldt, T., & Kinnune, U. (2006). The factor structure and factorial invariance of the 12-item General Health Questionnaire (GHQ-12) across time: Evidence from two community–Based samples. *Psychological Assessment*, *18*(4), 444–451.
Maslach, C., & Jackson, S. E. (1981). The measurement of experienced burnout. *Journal of Organizational Behavior*, *2*(2), 99–113.
Maslach, C., Schaufeli, W. B., & Leitner, M. P. (2001). Job burnout. *Annual Review of Psychology*, *52*(1), 397–422.
McCarthy, C. J., Lineback, S., & Reiser, J. (2015). Teacher stress, emotion, and classroom management. In E. T. Emmer & E. J. Sabornie (Eds.), *Handbook of classroom management* (pp. 301–321). New York: Routledge.
Midthassel, U. V., Bru, E., & Idsøe, T. (2000). The principal's role in promoting school development activity in Norwegian compulsory schools. *School Leadership & Management*, *20*(2), 147–160.
Ministry of Education. (2011) *Meld.St. 22 (2010–2011) Motivasjon – Mestring – Muligheter: Ungdomstrinnet [Motivation – mastery – possibilities]*.
Ministry of Education. (2016). *Fag – fordypning – forståelse — en fornyelse av kunnskapsløftet* [Subjects – Depth – Understanding – Renewing the knowledge promotion] Meld. St. 28 (2015–2016) Retrieved from https://www.regjeringen.no/no/dokumenter/meld.-st.-28-20152016/id2483955/sec1
Muijs, D. (2006). New directions for school effectiveness research: Towards school effectiveness without schools. *Journal of Educational Change*, *7*(3), 141–160. doi:10.1007/s10833-006-0002-7
Munthe, E. (2001). Professional uncertainty/certainty: How (uncertain) are teachers, what are they (un)certain about and how is (un)certainty related to age, experience gender, qualifications and school type? *European Journal of Teacher Education*, *24*(3), 355e368.
Munthe, E. (2003). Teachers' workplace and professional certainty. *Teaching and Teacher Education: an International Journal of Research and Studies*, *19*(8), 801–813.
Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide* (8th ed.). Los Angeles, CA: Muthén & Muthén.
Nir, A. E., & Bogler, R. (2008). The antecedents of teacher satisfaction with professional development programs. *Teaching and Teacher Education: an International Journal of Research and Studies*, *24*(2), 377–386.
Nylund-Gibson, K., Katherine, E., & Masyn, K. E. (2016). Covariates and mixture modeling: Results of a simulation study exploring the impact of misspecified effects on class enumeration. *Structural Equation Modeling*, *23*(6), 782–797.
Opfer, V. D., & Pedder, D. (2011). Conceptualizing teacher professional learning. *Review of Educational Research*, *81*(3), 376–407.
Pas, E. T., Bradshaw, C. P., & Hershfeldt, P. A. (2012). Teacher- and school-level predictors of teacher efficacy and burnout: Identifying potential areas for support. *Joural of School Psychology.*, *51*(1), 129–145.
Peugh, J., & Fan, X. (2012). How well does growth mixture modeling identify heterogeneous growth trajectories? A simulation study examining GMM's performance characteristics. *Structural Equation Modeling: A Multidisciplinary Journal*, *19*(2), 204–226.
Pianta, R., Hamre, B. K., & Mintz, S. L. (2012). *Classroom Assessment Scoring System (CLASS): Secondary class; manual.* Charlottesville, VA: Teachstone.
Pianta, R. C. (2006). Classroom management and relationships between children and teachers: Implications for research and practice. In C. M. Evertson & C. S. Weinstein (Eds.), *Handbook of classroom management: Research, practice and contemporary issues* (pp. 685–706). Mahwah, NJ: Lawrence Erlbaum Associates.

Pianta, R. C., Hamre, B. K., & Allen, J. P. (2012). Teacher-student relationships and engagement: Conceptualizing, measuring, and improving the capacity of classroom interactions. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of student engagement* (pp. 365–386). New York: Springer Nature.

Pianta, R. C. (2016). Classroom processes and teacher-student interaction: Integrations with a developmental psyhopatological perseptive. In D. Cicetti (Ed.), *Developmental psychopathology. Risk, recilience and intervention* (Vol. 4, 3rd ed., pp. 770–814). Holboken, NJ: Wiley.

Postholm, M. B., Midthassel, U. V., & Nordahl, T. (2012). *Teoretisk bakgrunnsdokument for arbeid med klasseledelse på ungdomstrinnet* [Theoretical background for classroom management in lower secondary schools]. Oslo: The Norwegian Directorate for Education and Training. Retrieved from https://www.udir.no/globalassets/filer/klasseledelse/ungdomstrinnet_bakgrunnsdokument_klasseledelse_vedlegg_1.pdf

Reddy, L. A., Dudek, C. M., Fabiano, G. A., & Peters, S. (2015). Measuring teacher self-report on classroom practies. Construct validity and reliability of the classroom strategies scale –Teacher form. *School Psychology Quarterly*, *30*(4), 513–533.

Richter, D., Kunter, M., Klusmann, U., Lüdtke, O., & Baumert, J. (2011). Professional development across the teaching career: Teachers' uptake. *Teaching and Teacher Education: an International Journal of Research and Studies*, *27*(Jan), 116–126.

Rogers, E. M. (2002). Diffusion of preventive innovations. *Addictive Behaviors*, *27*(6), 989–993.

Rogers, E. M. (2003). *Diffusion of innovations (5. utg.)*. New York: Free Press.

Ruzek, E. A., Hafen, C. A., Allen, J. P., Gregory, A., Mikami, A. Y., & Pianta, R. C. (2016). How teacher emotional support motivates students: The mediating roles of perceived peer relatedness, autonomy support, and competence. *Learning and Instruction*, *42*, 95–103.

Sandilos, L. E., & DiPerna, J. C.; The Family Life Project Key Investigators. (2014). Measuring quality in kindergarten classrooms: Structural analysis of the Classroom Assessment Scoring System (CLASS K–3). *Early Education and Development*, *25*(6), 894–914.

Schmidt, J., Klusmann, U., Lüdtke, O., Möller, J., & Kunter, M. (2016). What makes good and bad days for beginning teachers? A diary study on daily uplifts and hassles. *Contemporary Educational Psychology*, *48*(January), 85–97.

Shulman, L. S., & Shulman, J. H. (2004). How and what teachers learn: A shifting perspective. *Journal of Curriculum Studies*, *36*(2), 257–271.

Simmons, J. (2011). To improve schools, stop guessing and start using research. *Education Digest.*, *76*(9), 37–39.

Sinclair, C., Dowson, M., & McInerney, D. M. (2006). Motivations to teach: Psychometric perspectives across the first semester of teacher education. *Teachers College Record*, *108*(6), 1132–1154.

Skaalvik, E. M., & Skaalvik, S. (2011). Teacher job satisfaction and motivation to leave the teaching profession: Relations with school context, feeling of belonging, and emotional exhaustion. *Teaching and Teacher Education: an International Journal of Research and Studies*, *27*(6), 1029–1038.

Sørlie, M.-A., & Ogden, T. (2007). Immediate impacts of PALS: A school-wide multi-level programme targeting behaviour problems in elementary school. *Scandinavian Journal of Educational Research*, *51*(5), 471–492.

Spilt, J. L., Koomen, H. M., & Thijs, J. T. (2011). Teacher wellbeing: The importance of teacher-student relationships. *Educational Psychology Review*, *23*(4), 457–477.

Starnaman, S. M., & Miller, K. I. (1992). A test of a causal model of communication and burnout in the teaching profession. *Communication Education*, *41*(1), 40–53.

Tschannen-Moran, M., & Woolfolk Hoy, A. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education: an International Journal of Research and Studies*, *17*(7), 783–805.

Vangrieken, K., Meredith, C., Packer, T., & Kyndt, E. (2017). Teacher communities as a context for professional development: A systematic review. *Teaching and Teacher Education: an International Journal of Research and Studies.*, *61*(1), 47–59.

Vanhoof, J., Van Petegem, P., Verhoeven, J. C., & Buvens, I. (2009). Linking the policymaking capacities of schools and the quality of school self-evaluations: The view of school leaders. *Educational Management Administration & Leadership*, *37*, 667–686.

Vermunt, J. (2014). Teacher learning and professional development. In S. Krolak-Schwerdt, S. Glock, & M. Böhmer (Eds.), *Teachers' professional development* (pp. 79–95). Rotterdam: Sense Publishers.

Vermunt, J. D. (2011). Pattern in student learning and teacher learning. Similarities and differences. In S. Rayner & E. Cool (Eds.), *Style differences in cognition, learning and management* (pp. 173–187). New York: Routledge.

Virtanen, T., Pakarinen, E., Lerkkanen, M.-K., Poikkeus, A.-M., Siekkinen, M., & Nurmi, J.-E. (2017). A validation study of classroom assessment scoring system secondary in the finnish school context. *Journal of Early Adolescence*, 849–880. doi:10.1177/0272431617699944

Virtanen, T., Vaaland, G. S., & Ertesvåg, S. K. (2019). Associations between patterns of observed classroom interactions and teacher well-being. *Teaching and Teacher Education*, *77*, 240–252.

Wagner, W., Göllner, R., Helmke, A., Trautwein, U., & Lüdtke, O. (2013). Construct validity of student perceptions of instructional quality is high, but not perfect: Dimensionality and generalizability of domain-independent assessments. *Learning and Instruction*, *28*, 1–11.

Wagner, W., Göllner, R., Werth, S., Voss, T., Schmitz, B., & Trautwein, U. (2016). Student and teacher ratings of instructional quality: Consistency of ratings over time, agreement, and predictive power. *Journal of Educational Psychology*, *108*(5), 705–721.

Wayne, A. J., Yoon, K. S., Zhu, P., Cronen, S., & Garet, M. S. (2008). Experimenting with teacher professional development: Motives and methods. *Educational Researcher*, *37*(8), 469–479.

Westergård, E. (2007). Do teachers recognise complaints from parents, and if not, why not? *Evaluation & Research in Education*, *20*(3), 159–178.

Westergård, E., Ertesvåg, S. K., & Rafaelsen, F. (2019). Preliminary validity of the classroom assessment scoring system (CLASS) in Norwegian lower secondary schools. *Scandinavian Journal of Educational Research*, *63*(4), 566–584.

Wickrama, K. A. S., Lee, T. K., O'Neal, C. W., & Lorenz, F. O. (2016). *Higher-order growth curves and mixture modeling with Mplus: A practical guide*. New York, NY: Routledge.

Wubbels, T., Brekelmans, M., Den Brok, P., Wijsman, L., Mainhart, T., & van Tartwijk, J. (2015). Teacher-student relationships and classroom management. In E. T. Emmer & E. J. Sabornie (Eds.), *Handbook of classroom management* (2nd ed., pp. 363–386). New York: Routledge.

Wubbels, T., & Brekelmans, M. (2005). Two decades of research on teacher-student relationships in class. *International Journal of Educational Research*, *43*(1–2), 6–24.