University of Groningen

A longitudinal study of induction on the acceleration of growth in teaching quality of beginning teachers through the eyes of their students

Maulana, Ridwan; Helms-Lorenz, Michelle; van de Grift, Wim

Published in:
Teaching and Teacher Education

DOI:
[10.1016/j.tate.2015.07.003](https://doi.org/10.1016/j.tate.2015.07.003)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2015

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Maulana, R., Helms-Lorenz, M., & van de Grift, W. (2015). A longitudinal study of induction on the acceleration of growth in teaching quality of beginning teachers through the eyes of their students. *Teaching and Teacher Education*, 51, 225-245. <https://doi.org/10.1016/j.tate.2015.07.003>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

A longitudinal study of induction on the acceleration of growth in teaching quality of beginning teachers through the eyes of their students

Ridwan Maulana*, Michelle Helms-Lorenz, Wim van de Grift

Department of Teacher Education, University of Groningen, The Netherlands

HIGHLIGHTS

- We examine changes in teaching quality during the first three years of teaching.
- We explore personal and contextual factors affecting perceived changes in teaching quality.
- Teaching quality increases during the first three years of teaching.
- Changes in teaching quality differ between beginning teachers.
- Certification, gender, and induction programs explain changes in perceived teaching quality.

ARTICLE INFO

Article history:

Received 22 September 2014

Received in revised form

5 June 2015

Accepted 7 July 2015

Available online 6 August 2015

Keywords:

Teaching quality

Beginning teachers

Multilevel growth curve modelling

Student perceptions

Secondary education

ABSTRACT

This study examines the development of beginning teachers' teaching quality during the first three years of teaching. Schools were randomly assigned to an experimental condition, to develop support programs (induction) for beginning teachers. Induction focused on personal coaching and the enhancement teacher autonomy and connectedness. Analyses were based on a total sample of 276 beginning teachers and 4932 students from 63 schools over the course of three years. Students perceived more rapid teaching quality increases over time in the experimental condition compared to the control condition. Certification status, gender, and induction programs explain differences (and changes) in perceived teaching quality.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

The improvement of students' learning and outcomes is and continues to be a global educational concern. Teachers, who are seen as central figures in promoting students' academic performance to the highest level possible, are and have been a major focus of educational research. Research has consistently shown that a significant factor for lessening students' behaviour problems and improving their academic gains is through the development of teaching quality (Alderman & Nix, 1997; Darling-Hammond, Chung, & Frelow, 2002; Kyriakides, Creemers, & Antoniou, 2009).

This suggests that students can benefit more from having effective teachers. When students are taught by less effective teachers, they may not reach their full potential in learning.

In everyday practice, schools employ both experienced as well as beginning teachers in order to meet the demand of teaching staff. Unfortunately, research indicates a general gap in teaching quality between experienced and beginning teachers. Generally, teaching quality of beginning teachers is lower compared to that of experienced teachers (Maulana, Opdenakker, Stroet, & Bosker, 2012). Beginning teachers tend to face difficulties in providing a supportive learning climate, managing the classroom effectively, displaying clear instruction, activating students' learning, attending to differences between students' needs, and teaching students how to learn effectively (Maulana et al., 2012; Maulana, Helms-Lorenz, & Van de Grift, 2015). Beginning teachers lack the conceptual

* Corresponding author. Department of Teacher Education, Grote Kruisstraat 2/1, 9712 TS Groningen.

E-mail address: r.maulana@rug.nl (R. Maulana).

understanding and the abilities to extend the lessons beyond the basics (Wilson, Floden, & Ferrini-Mundy, 2002). Hence, one may expect that the behaviour of beginning teachers may not be as effective as that of experienced teachers. Indeed, studies in the United States reveal that experienced teachers produce higher student achievement compared to inexperienced teachers (Greenwald, Hedges, & Lane, 1996; Rivkin, Hanushek, & Kain, 2005).

However, less is known about how beginning teachers' teaching quality develops over time. From the perspective of teacher education and in-service teacher professional development programs, it is important to understand more about how beginning teachers' teaching quality develops and which teaching skills to support to facilitate development over time. If changes in teaching quality differ between beginning teachers, then this will inform school mentors and teacher educators regarding the need to attend to individual beginning teacher differences in their professional trajectory over time. Furthermore, understanding changes in teaching quality could give researchers clues regarding the peak moment of beginning teachers' teaching quality so effective adjustments regarding interventions for improving teaching behaviour can be made. This suggests that the nature of interventions aiming at the improvement of teaching quality may need to be adjusted over time. Additionally, it is inconclusive which factors contribute to explaining differences and changes in beginning teachers' teaching quality. Undoubtedly, knowing significant factors for (the development of) beginning teachers' teaching quality is useful information for teacher education and in-service teacher professional development program.

In order to extend the knowledge base of the development of teachers' teaching quality, the aims of the present study were to investigate variability and change, and the effects of personal and contextual factors (particularly induction programs to support teachers' professional development in teaching) on teaching quality during the first three years of professional teaching.

2. Theoretical framework

2.1. Teaching quality

Teaching quality covers some or all aspects of teacher-student relationships in the classroom (Cornelius-White, 2007). Teaching and teacher effectiveness and learning environments research, which serve as theoretical frameworks for studying teaching quality in the current study, have shown that the quality of teaching practices are significant predictors of students' learning and outcomes (Creemers & Kyriakides, 2008; Hattie, 2012). In contemporary conceptualizations of teaching and learning processes, both cognitive and motivational aspects are taken into account.

Ko, Sammons, and Bakkm (2013) present a comprehensive overview of definitions of effective teaching ranging from narrow to broad encompassing definitions. Our operationalization of effective teacher behaviour corresponds to their first definition: The effectiveness of observable behaviours seen during classroom observation of a typical lesson. This aspect of teacher effectiveness is the core of all the mentioned broader definitions. The current study focusses on the effects of induction support on observable behaviour in the classroom of beginning teachers. For this goal, the least broad definition suffices. We intend to measure the behavioural effects of support in the classroom, therefore our operationalization of teaching quality is narrower than the broader definition of effective teaching we support, even though we fully acknowledge the scope of the above mentioned aspects of effectiveness of teaching and ways to measure it.

Evidence-based teaching practices have their historical roots in

learning theories, put forward and studied for many more decades than studies concerning effective teacher behaviour. Learning theories (behaviorism, cognitive theories and constructivist¹ theories) stem from different (ancient) epistemological viewpoints (Relativism, Objectivism) defending the world as either *constructed* or as *given* respectively (Yeaman, Hlynka, Anderson, Damarin, & Muffoletto, 1996). The instructional principles derived from the learning theories lead to different manifestations of instruction. Superficially, rival theories seem to lead to opposing practices. Recent research has been triggered by these differences in practice, leading to questions as: what works best, and under which circumstances do certain teacher practices lead to higher student outcomes (teaching effectiveness)? The meta-analysis of Hattie (2009) revealed that teaching practices stemming from diverse learning theories are effective (e.g., mastery learning of the behaviouristic learning theory and meta-cognitive strategies derived from the cognitivist learning theory and problem-solving teaching derived from the constructivist approach). This evidence adds to the validity of the learning theories and might stimulate the refinement and integration of the valuable insights of these theories in the future. This might give way to a higher order eclectic learning theory accommodating more than one learning theory.

Research suggest that social contexts such as school and departmental leadership/culture (Little, 1982), characteristics of school population and changes in social conditions (Hutchings, Menter, Ross, Thomson, & Bedford, 2000), and professional development opportunity (Day, 1999) affect teaching quality. Furthermore, teacher intrapersonal aspects play a significant role for teachers to achieve, maintain, and improve teaching practices. Those aspects include motivation and morale (Troman & Woods, 2001), passion for teaching (Day, 2004), adaptability for change (Fullan & Hargreaves, 1992), age and developmental phase (Fessler, 1995), self-efficacy (Rosenholtz, 1989), and satisfaction through students' progress and achievement (Tschannen-Moran, Hoy, & Hoy, 1998). In short, social environments and teacher intrapersonal aspects are important factors determining teacher effectiveness.

2.2. Evidence-based teaching practice domains

Based on reviews of evidence-based effective teachers' behaviour research, Van de Grift (2007; Maulana, Opendakker, Stroet, and Bosker, 2013) identified six observable domains of teaching quality. These domains include safe and stimulating learning climate, efficient classroom management, clarity of instruction, activating learning, adaptation to students' learning needs, and teaching learning strategies. The six domains of teaching quality mentioned above are also in accordance with other models and empirical findings (Danielson, 2013; Ko et al., 2013; Pianta & Hamre, 2009). Five of the 10 effective behaviour domains found by Ko et al. (2013) coincide with the 6 above mentioned domains.

Safe and stimulating learning climates refer to aspects of teacher-student relationships such as creating a safe and relaxing classroom climate, showing respect to students and ensuring that students respect the teacher and their peers, and encouraging self-confidence of students (Maulana et al., 2012; Cornelius-White, 2007; Hattie & Clinton, 2008; Smith, Baker, Hattie, & Bond, 2008; Teodorovic, 2011; Willms & Somer, 2001). Effective classroom management is another important domain of teaching quality.

¹ It goes beyond our scope to discuss the status and quality of the learning theories, rather our aim is to argue that different lines of thought, concerning how we learn, lead to different forms of instruction. These different forms of instruction have proven to be more or less effective.

Research indicates that efficient classroom management is a strong predictor of students' learning and outcomes (Maulana et al., 2012; Carnine, Dixon, & Silbert, 1998; Houtveen, Booij, De Jong, & Van de Grift, 1999; Scheerens & Bosker, 1997). Several indicators of teachers' teaching quality associated with efficient classroom management include ensuring that the lesson begins and ends on time, managing lesson transition efficiently, minimizing time for task-unrelated matters, dealing with students' misbehaviour efficiently, preparing the lesson well, and managing an efficient lesson structure (Maulana et al., 2012; Marzano, 2003; Opdenakker & Minnaert, 2011; Wang, Reynolds, & Walberg, 1995; Yair, 2000). Clarity of instruction covers aspects of instructional quality that are important for students' learning such as clear lesson structure, good interchange of explanations and lesson presentations, management of independent work, and clear assignment of individual and group works (Kindsvatter, Wilen, & Ishler, 1988; Mortimore, Sammons, Stoll, Lewis, & Ecob, 1988; Rosenshine, 1980). Additionally, checking frequently that pupils understand the learning material is also important (Hattie & Clinton, 2008; Kame'enui & Carnine, 1998; Pearson & Fielding, 1991; Rosenshine & Meister, 1997; Smith et al., 2008).

Activating learning integrates behaviour that optimizes learning outcomes such as promotion of active learning, intensification of instructions and avoiding excessive work seats (Hampton & Reiser, 2004; Lang & Kersting, 2007). Other important aspects of activating learning for productive learning outcomes are activation of prior knowledge, making use of "advance organizers", and making sure that pupils are aware of the relevance of the lesson content (Nunes & Bryant, 1996; Pressley et al., 1992). Activating learning is related to the quality of teacher-student and peer interactions (Meeuwisse, Severiens, & Born, 2010). Subsequently, the quality of teacher-student interactions is related to the improvement of motivational outcomes (Maulana et al., 2013; Opdenakker, Maulana, & den Brok, 2012). Adaptation in teaching means that teachers recognize the characteristics of their students they teach and understand that different students have learning needs in order to progress in learning. Subsequently, teachers need to adapt and adjust their teaching to the learning needs of their students. Teachers should emphasize the needs of each individual student when displaying instructional methods. In the literature of differentiation, the term "differentiated instruction" is recognized. This term refers to "a philosophy of teaching purporting that students learn best when their teachers effectively address variance in students' readiness levels, interests, and learning profile preferences" (Tomlinson, 2005, p. 263). The main goal of differentiated instruction is to maximize the learning potential of each student (Tomlinson, 2003). Research showed that differentiated instruction leads to better learning performance (Koeze, 2007; Reis, McCoach, Little, Muller, & Kaniskan, 2011). Several indicators of teachers' teaching quality associated with differentiation in teaching include devoting extra time and additional instructions, pre-teaching and re-teaching, and implementing various effective teaching methods (Houtveen et al., 1999; Kindsvatter et al., 1988; Lundberg & Linnakylä, 1992; Pearson & Fielding, 1991; Sijtsma, 1997).

Finally, teaching learning strategies is "a heuristic that serves to support students, facilitating the development of internal procedures that enable them to perform the higher-level procedures (Van de Grift, 2007). Teaching learning strategies suggest that teachers support the use of metacognitive strategies in their classroom teaching. Metacognitive strategies provide a framework to help students achieve higher level of learning skills (Carnine et al., 1998). One way of supporting the use of metacognitive strategies is scaffolding. Scaffolding is a form of temporary support provided by teachers (or by peer students) that function as a bridge between students' existing and desired skills. Research shows that

teachers who display modelling explicitly, deliver scaffolding and provide corrective feedback, contribute to the learning performance of their students significantly (Good & Brophy, 1989; Hattie & Clinton, 2008; Houtveen & Van de Grift, 2007; Rosenshine & Stevens, 1986; Slavin, 1996; Smith et al., 2008).

2.3. Development of teaching quality

In general, the growth of teachers' teaching quality can follow three possible developmental trends: Increasing, stabilizing (level-off), and declining. The combination of the mentioned three trends (curvilinear) is possible too. The first and the last developmental trends can be characterized by a linear curve (no acceleration or deceleration in growth), or quadratic and cubic growth curves (with acceleration or deceleration in growth). Acceleration and deceleration is related to the steepness in and the direction of the growth over time. Patterns of change can differ between teachers (interpersonal differences) and within teacher (intrapersonal differences) over time (Nesselrode, 2001). Interpersonal change over time indicates variations in the shape and pacing of change from one teacher to the other, while intrapersonal change over time indicates some irregularity of change within the teacher across situations (Malmberg, Hagger, Burn, Mutton, & Colls, 2010).

There is emerging research that focuses on changes in teaching quality over time. A study combining surveys and teacher interviews in the United States shows that professional activities deepen and grow as teaching experience increases (Shagrir, 2012). Similarly, a study based on self-reported teaching quality in Singapore reveals that beginning teachers reported a significant increase in lesson planning, instructional support and classroom management over time (Choy, Wong, Lim, & Chong, 2013). Another research based on student surveys from Indonesia shows that teaching quality, studied from teacher-student interpersonal perspective, increases across the school year (Maulana et al., 2013). In England, a longitudinal observation study on teachers from the pre-service to two years professional practice shows that the development of teaching quality can follow a more complex pattern than linear and, depending on the domain of teaching quality (Malmberg et al., 2010). They found that classroom management increases linearly over time, while emotional and instructional support increases during pre-service education and the first year of teaching before it decreases subsequently. A similar trend of findings is also evident from the Netherlands. Based on student surveys, it was found that teachers' quality in maintaining interpersonal control and cooperativeness with their students decreases over time, but the decrease in teachers' interpersonal cooperativeness is less steep toward the end of the school year (Opdenakker et al., 2012).

Furthermore, observation research comparing the quality of teachers' interpersonal relationships with students in The Netherlands and Indonesia across the school year show that teachers' involvement decreases during the first four months and continue to increase until the eight month, then decreases again until the end of the school year (Maulana et al., 2013). Research suggests that intra-personal variability in the development of teaching quality can be perceived as teacher adaptation to or coping with situational demands (Frenzel & Götz, 2007; Gibson & Dembo, 1984). However, changes in teaching quality over time are not always visible. Based on observation studies, findings from Germany reveal that teaching quality is stable across three lessons (Kunter & Baumert, 2006) as well as across six lessons (Seidel, et al., 2002). Consistent with this, an observation study from the Netherlands and Indonesia indicates that teachers' devotion to elements of lesson structure tends to be stable across the school year (Maulana et al., 2012).

Taken together, the knowledge about changes in teachers' teaching quality remains incomplete. To our knowledge, there are no studies examining changes in beginning teachers' teaching quality by taking into account the hierarchical structure of the data, by paying attention to differences between teachers and by studying changes in teaching practice domains over time across three years of professional practice. The present study aims to contribute to filling this gap by studying teaching from the perspective of students.

2.4. Students' perceptions of teaching quality

One way of measuring teachers' teaching quality is by using an observation instrument, which is considered as the most objective tool to tap the actual teachers' teaching practices (Lasagabaster & Sierra, 2011; Worthen, Sanders, & Fitzpatrick, 1997). However, there is a number of reasons that studying teaching from the perspective of students contributes to the understanding of teachers' teaching practices in a complementary way. Student perceptions are based on their (classroom) experiences over time, not merely from a single snapshot or limited number of observations. Therefore, what students perceive may actually be more important than what observers would observe since student perceptions steer individual learning behaviour, based on own judgements and insights. Indeed, research suggests that whether students will academically engage in and successfully complete the expected activities depends on their perceptions of the quality and integrity of teachers' behaviour (Brok, Bergen, Stahl, & Brekelmans, 2004). Similarly, other studies show that student perceptions are mostly more predictive of various student outcomes (i.e., motivation, engagement, achievement) compared to observations (De Jong & Westerhof, 2001; Seidel & Shavelson, 2007) and teachers' subjective perceptions of their own teaching behaviour (Maulana, Opdenakker, den Brok, & Bosker, 2011; Scantlebury, Boone, Kahle, & Fraser, 2001). From the economic point of view, the use of questionnaires is mostly preferred because it is considered as the most cost-effective research tool compared to observations (Fraser, 1991).

2.5. Teacher induction

Beginning teacher attrition and the transition shock from teacher education to the teaching profession have been recurring problems over the last decades in many countries. Teacher induction programs have been developed to solve these problems and can be defined as a more or less planned and formalized arrangement in which beginning teachers are assisted and supported in becoming competent and effective professionals.

As Feiman-Nemser, Schwille, Carver, and Yusko (1999) put it on p 20 "a program suggests something formal and deliberate. It invokes such descriptive analytic categories as goals, curriculum, organizational structure, staff, clients, funding and evaluation. Yet, deciding what counts as an induction program is not as straightforward as the term implies". The term induction program suggests a fixed, prescriptive curriculum. A curriculum specifies topics to be understood and specifies the standard to be achieved. We prefer the term tailor-made arrangement, implying flexibility in depth and pace. The degree of resilience and perceived self-efficacy of the beginning teacher (BT) in relation to the developmental goals to be met, should be the leading principles in determining the degree of intensity (frequency) and the kind of support to be provided.

This study evaluates a specific goal of the nationwide experiment in The Netherlands: supporting the growth of effective teaching behaviour during the first three years in the teaching profession.

3. The current study

3.1. Dutch secondary education

After finishing primary education (age 12–13), students can enter one of the following secondary education tracks: (1) Pre-vocational secondary education (VMBO, 4 years), (2) Senior general secondary education (HAVO, 5 years), and pre-university education (VWO, 6 years). The enrolment is based on students' performance in the standardized test at the end of primary education and their teachers' recommendation. School choice is open and parents are obliged to choose a school. School careers of students are highly differentiated in secondary education in terms of curricula, prestige, number of years' duration, final qualifications, and possible destination after secondary education (Sykes & Kuyper, 2013).

Generally, schools offer multiple tracks. In the first year, students can follow a program which does not differentiate yet between the two tracks or more. In the second year, students will be assigned to the definitive track based on their performance in the first year. Social class and ethnic group are not distributed evenly across the different tracks in secondary education. The lower track is over-represented by students of parents with lower levels of education, while the higher tracks are under-represented by students from lower educational backgrounds (Herweijer, 2009). Regardless of this discrepancy, research suggest that ethnic composition is not related to achievement in secondary education, after taking into account socio-economic composition. School composition accounts for little, if any, of the observed differences in achievement across ethnic/social background (Sykes & Kuyper, 2013). There is a teacher shortage in The Netherlands (Maulana et al., 2015). This implies that the teaching profession in the country cannot be considered as highly competitive. There are various steps taken to bring back ex-teachers to the profession and to recruit new and highly qualified teachers to keep up with demand.

3.2. Teacher induction program in The Netherlands

In the experimental group, an induction arrangement was developed and implemented in cooperation with the researchers, institute-based educators and school-based educators for three consecutive school years. The intervention aims to support teacher autonomy, competence and relatedness as described in the self-determination theory of Deci and Ryan (2000). These fundamental needs fuel teacher's development and specify the required conditions for psychological and professional growth, integrity, and well-being. Teacher autonomy is supported by taking the individual teacher's concerns into account regarding the conditions needed to thrive. This can be achieved by (1) considering the need to reduce (or increase) the workload. Teacher autonomy is nourished by increasing the competence level, by (2) supporting effective teaching behaviour in the classroom (this is enhanced by working together with colleagues when preparing lessons, observing each other etc) and by (3) stimulating professional development (developing transparent induction goals and assessment procedures, clear human resource management and policy). The need for relatedness calls for (4) facilitating activities to get familiar with the own students, school policy, school rules, and with colleagues. The development tasks proposed by Feiman-Nemser (2001) were used as leading principle to streamline activities into a coherent arrangement encompassing three years. These principles are: (1) gaining local knowledge of students, curriculum, and school context, (2) designing responsive curriculum and instruction, (3) enacting a beginning repertoire in purposeful ways, (4) developing a professional identity and, (5) learning in and from practice. The

guidelines provided by [Ingersoll and Smith \(2004\)](#) were used as framework for the development of induction arrangements: the arrangement has a duration of three years because the greatest drop-out occurs in the first five years on the job. BT's are to be provided with a mentor, ideally from the same subject field and available on a daily basis for timely support, opportunities for common planning time with other teachers in the same subject, opportunities for regular scheduled collaboration with other teachers, opportunities for being part of an external network of teachers, opportunities for structured peer coaching, opportunities for classroom observations with feedback, workshops, opportunities to work with professional development plans, opportunity to work with portfolio's, many activities because as the number of components in the induction arrangements increases, the probability of retention increases. Workload reduction was added to the arrangements as many studies reveal that BTs need extra time to reflect, collaborate and to observe others (e.g., [Algozzine, Gretes, Queen, & Cowan-Hathcock, 2007](#); [Lazovsky & Reichenberg, 2006](#); [Nielsen, Barry, & Addison, 2006](#)). Guidelines of the national induction program are provided in [Appendix 2](#) and [3](#). [Appendix 2](#) specifies working condition features that can support the autonomy of the beginning teacher. [Appendix 3](#) specifies induction domains that can support the relatedness to the school (school organization and context) and the students (most domains), domains to support autonomy (learning of the own education practice, professional development) and domains to support competence (effective teaching behaviour).

3.3. Research questions

In the present study, the following questions will be answered:

1. What is the general perception of students concerning their beginning teachers' teaching quality during the first three years of professional practice?
2. What are differences in perceived beginning teachers' teaching quality, between schools, between teachers, and across the first three years of professional practice?
3. Does student perception of beginning teachers' teaching quality change during the first three years of professional practice?
4. Are there differences between individual beginning teachers with regard to change across 3 years as perceived by students?
5. How do contextual features (class size, grade level, induction program) and beginning teacher characteristics (teaching subject, gender, certification) explain differences and changes in beginning teachers' perceived teaching quality?

4. Method

4.1. Sample and procedure

The data was drawn from a cohort of a longitudinal study, the national project of induction arrangements in secondary education in The Netherlands. During the academic years 2010 through 2013, beginning teachers were followed during the first three years of their professional practice and their students were asked to fill in the questionnaire on teaching quality voluntarily. Agreement between researchers, schools, and beginning teachers was established prior to the execution of the study. In the non-induction group condition (control), beginning teachers followed regular arrangement in their schools and did not receive any guidance. Treatment

and control groups were assigned randomly after school recruitment (at the school level). About half of the participating teachers received small incentives.²

In this longitudinal project, a total of 338 beginning teachers and their 6770 students from 68 schools participated in the study. Due to technical and practical reasons typically encountered in longitudinal studies, the present study included the data of 276 beginning teachers and their 4932 students from 63 schools (first wave). In the second wave, 29% of the teachers dropped-out in the study, and an additional 13% of the teachers dropped-out in the third wave. Out of 276 teachers, 48% participated in all waves. Teachers who dropped in the first wave did not participate in the subsequent waves.

Of the beginning teachers, 56% were female, 25% taught math and science subjects, 64% taught first and second grades of secondary education, 74% was certified teachers, 74% taught large class size, and 53% followed induction programs in their school. Class size ranged between 10 and 37 students. Across three years, a questionnaire of teaching quality was distributed to students of the corresponding beginning teachers. Regular classrooms were selected by the beginning teachers and mentors to avoid outliers (choosing extreme groups of students). Students filled in the questionnaire at the end of each school year.

4.2. Measures

4.2.1. Teaching quality

To capture students' perceptions of beginning teachers' teaching quality, we used an existing questionnaire on teaching quality ([Maulana et al., 2015](#)). This questionnaire was based on a standardized observation instrument on teaching quality ([Maulana et al., 2013](#); [Maulana et al., 2015](#)). The questionnaire measures six domains of teaching quality in the 24 items configuration: learning climate (3 items, example: "My teacher treats me with respect"), classroom management (3 items, example: "My teacher applies clear rules"), clarity of instruction (5 items, "My teacher explains everything clearly to me"), activating learning (3 items, example: "My teacher involves me in the lesson"), adaptation (7 items, example: "My teacher connects to what I know or am capable of"), and teaching learning strategies (3 items, example: "My teacher explains how I should study something"). All items were provided on a 4-point Likert scale ranging from 1 (Completely disagree) to 4 (Completely agree). See [Appendix 1](#) for the student questionnaire.

The questionnaire has been reported to be a valid and reliable measure of teaching quality ([Maulana et al., 2015](#)). For the present sample, we analyzed reliability and validity of the measure in a number of ways. Regarding the reliability of items over measurement occasions, results show that the internal consistency of all scales are above the satisfactory level, ranging from 0.66 (learning climate) to 0.87 (Adaptation) at the student level and from 0.86 (learning climate) to 0.98 (adaptation) at the aggregated class level (see [Table 1](#)).³ Moreover, mean inter-scale correlations range from 0.58 (Classroom management) to 0.75 (adaptation), indicating that although there is an overlap among domains of teaching quality, the scales could measure distinct aspects of teaching behaviour. According to previous research on the same measure ([Maulana et al., 2015](#); [Van de Grift, 2007](#); [Maulana et al., 2013](#)), the six domains forms a unidimensional construct called teaching quality. We performed confirmatory factor analysis to validate the unidimensional factor pattern with the present data (using Mplus,

² Beginning teachers received vouchers with a value of 15 dollars in year 1, 20 dollars in year 2 and 30 dollars in year 3 for participating in the data collection.

³ Level-1 Cronbach's alpha reliability was computed using student level data, while level-2 Cronbach's alpha reliability was computed using aggregated class level data.

Table 1
Reliability and descriptive statistics of teaching quality measure.

Domain	Internal consistency						Mean and SD						Mean inter-scale correlation		
	Time 1		Time 2		Time 3		Time 1		Time 2		Time 3		Time 1	Time 2	Time 3
	Student	Class	Student	Class	Student	Class	Mean	SD	Mean	SD	Mean	SD			
1. Learning climate	0.70	0.92	0.66	0.89	0.66	0.86	3.10	0.63	3.23	0.55	3.32	0.50	0.65	0.59	0.59
2. Classroom management	0.74	0.95	0.68	0.94	0.70	0.93	2.79	0.71	3.00	0.61	3.10	0.56	0.67	0.59	0.58
3. Clear instruction	0.83	0.95	0.77	0.91	0.77	0.89	3.01	0.65	3.18	0.55	3.26	0.50	0.72	0.65	0.64
4. Activating learning	0.76	0.96	0.70	0.91	0.70	0.90	2.78	0.71	2.96	0.61	3.05	0.56	0.71	0.64	0.63
5. Adaptation	0.87	0.98	0.84	0.96	0.84	0.96	2.76	0.63	2.91	0.54	3.00	0.51	0.75	0.67	0.67
6. Teaching strategies	0.72	0.92	0.67	0.88	0.67	0.86	2.77	0.69	2.94	0.60	3.02	0.56	0.67	0.62	0.62

Muthén & Muthén, 1998–2012). According to the fit statistics criteria proposed by Hu and Bentler (1999), results showed that the unidimensional factor of teaching quality was deemed acceptable ($\chi^2_{df=235} = 480.55$, $p < .01$, CFI = 0.96, TLI = 0.95, RMSEA = 0.06, SRMR = 0.06).⁴

4.2.2. Contextual and teacher characteristics

Together with the questionnaire on teaching quality, we included several contextual and teacher characteristics to investigate whether or not these characteristics matter for teaching quality of beginning teachers. The following contextual features were explored: grade level (0 = lower grade (1st and 2nd grade), 1 = higher grade (3rd–6th grade)), class size (0 = small class (≤ 20 students), 1 = large class (> 20 students)), and beginning teacher group condition (0 = control group, 1 = induction group). With regard to teacher characteristics, the following variables were examined: teacher gender (0 = male, 1 = female), teaching subject (0 = science (math and science-related subjects: physics, biology, and chemistry)), 1 = non science (e.g., economy, history, English, philosophy), and teacher certification status (0 = uncertified, 1 = certified). Within Dutch schools, teachers are viewed by their students as certified teachers irrespective of their certification status. Students in The Netherlands cannot see the difference between a certified and an uncertified teacher in terms of responsibility and status of the teacher. Because of teacher shortages uncertified teachers are given full responsibility for classes and are obliged to achieve certification within 2 years. Pre-service teachers in their last year of teacher training are also given full responsibility for their classrooms. The certified-as well as the uncertified beginning teachers have temporary contracts, are coached and lessons are observed by other teachers. Permanent appointments are considered at least one year after certification and after an additional assessment performed by the school leadership. Therefore schools as well as students view uncertified teachers as members of “beginning” staff, teachers on their way to permanent appointments.

4.3. Analytic strategy

The data of the present study were naturally and systematically structured in a hierarchical order (i.e., measurement occasions nested within beginning teachers, beginning teachers nested within classes). With this data structure, the application of traditional (single-level) regression analyses is less appropriate because the results produced will tend to be overestimated. This is because

ignoring the hierarchical structure of the data results in less accurate fixed effect and standard error estimates and inflation of lower level variance estimates (Chen, 2012; Snijders & Bosker, 2012). Multilevel modelling is more appropriate for this type of data instead (Snijders & Bosker, 2012). When the data are ordered hierarchically and longitudinally, a more complex multilevel modelling, called multilevel growth curve modelling (MGCM), is the most appropriate. With this method, not only the hierarchical structure of the data is taken into account, but also the multiple measurements over time of teaching quality and predictor variables. MGCM is an extension of mixed-effect regression model (MRM) applied to multilevel and longitudinal data (see Rasbash, Steele, Browne, & Goldstein, 2014 for more detail). This model is “more flexible in term of repeated measures and does not require restrictive assumptions concerning missing data across time and the variance–covariance structure of the repeated measures. Hence, MRM can be used for unbalanced/incomplete longitudinal data” (Nakai & Ke, 2011, p. 4). Even if the missing pattern is not missing completely at random (MCAR), MGCM can handle the missing data and produce efficient estimates (Rasbash, Steel, Browne, & Goldstein, 2014).

To investigate the missing data pattern, we conducted diagnostic analyses involving Little's MCAR test (Hill, 1997), descriptive statistics evaluations of missing versus non-missing data, and tabulated data patterns (see Appendix 4). The Little's MCAR test resulted in a chi-square = 73.49 (df = 54; $p = .04$). The p value is significant at 5% level, but not significant at 1% level. This indicates that the missing data may be missing completely at random (MCAR, i.e., no obvious systematic pattern exists to the missing data). Furthermore, the mean differences between missing and non-missing data when missingness in categorical variables teaching subject, class size, grade level, and certification are taken into account are minimal. The missingness of the mentioned background variables does not seem to affect the means of several of the scales of teaching quality. This can be an indication that the missingness pattern may be MCAR. Further inspection of the crosstabulations of categorical variables versus teaching quality scales show that the number of missing values in the teaching quality scales do not appear to vary much between teaching subject, class size, grade level, and certification categories (the overall discrepancy is lower than 10%). The tabulated missing patterns show that when cases are missing in the second wave of measurement, they are also missing in the third wave of measurement. This is not surprising because teachers who dropped out in the second wave did not participate in the third wave. Again, information from this missing pattern diagnosis shows that the mean of teaching quality scales do not vary considerably depending on the missing pattern.

To examine the first research question, we performed descriptive analyses for each domain of teaching quality across three years. To answer the remaining four research questions we conducted multilevel growth curve modelling (with MLwiN, Rasbash, Charlton, Browne, Healy, & Cameron, 2005). The modelling was

⁴ Fit indexes cutoff levels for determining model fit proposed by Hu and Bentler (1999): $\chi^2 p > .05$, CFI (Comparative Fit Index) ≥ 0.95 , TLI (Tucker–Lewis Index) ≥ 0.95 , RMSEA (Root Mean Square Error of Approximation) < 0.06 , SRMR (Standardized Root Mean Square Residual) ≤ 0.08 .

Table 2
Correlations between teaching behaviour variables, personal, and contextual characteristics.

	1	2	3	4	5	6	7	8	9	10	11	M	SD
1. Learning climate	–											3.21	0.58
2. Classroom management	0.58**	–										2.95	0.65
3. Clear instruction	0.67**	0.66**	–									3.14	0.58
4. Activating learning	0.61**	0.69**	0.68**	–								2.92	0.64
5. Adaptation	0.68**	0.66**	0.74**	0.74**	–							2.88	0.57
6. Teaching strategies	0.59**	0.61**	0.69**	0.66**	0.74**	–						2.90	0.63
7. Teaching subject ^a	0.03**	0.07**	0.04**	0.06**	0.04**	0.03**	–					75%	–
8. Class size ^b	0.03**	0.04**	0.06**	0.03**	0.01	0.02*	0.14**	–				75%	–
9. Condition ^c	–0.06**	–0.05**	–0.10**	–0.05**	–0.07**	–0.07**	–0.05**	–0.03**	–			53%	–
10. Grade level ^d	0.00	–0.05**	–0.01	–0.06**	–0.01	–0.04**	–0.07**	–0.15**	–0.08**	–		36%	–
11. Teacher gender ^e	–0.06**	0.00	–0.06**	–0.05**	–0.04**	–0.04**	0.09**	0.05**	0.01	–0.10**	–	56%	–
12. Certification ^f	0.26**	0.31**	0.27**	0.29**	0.26**	0.27**	0.19**	–0.03	–0.19**	–0.08*	0.22**	74%	–

Note. * $p < .5$, ** $p < .01$. ^a 0 = science, 1 = non-science, ^b 0 = small class, 1 = large class, ^c 0 = control group, 1 = induction group, ^d 0 = lower grade, 1 = higher grade, ^e 0 = male, 1 = female, ^f 0 = uncertified, 1 = certified.

Table 3
Distribution of the total variance of teaching behaviour over the school, teacher and occasion level (percentages).

	Learning climate	Classroom management	Clear instruction	Activating learning	Adaptation	Teaching strategies
School level	4%	10%	7%	9%	8%	9%
Teacher level	48%	44%	48%	47%	43%	42%
Occasion level	48%	46%	45%	44%	49%	49%

started by estimating a null model (Model 0) to answer the second research question. Furthermore, we included fixed effects of time (linear, quadratic; Model 1) to answer the third research question. Next, we added random effects of time (linear) to answer the fourth research questions (Model 2). Finally, we included the remaining predictors in the model to answer the fifth research question (Model 3).⁵ The modelling was done using a stepwise procedure and separately for each domain of teaching quality. Significant predictors at $p < .05$ were retained.⁶ The fixed effects in the model were tested by using t-ratio coefficients for a significant effect of a variable (Snijders & Bosker, 2012).

5. Results

5.1. General picture of teaching quality

Based on the total mean scores across the first three years of professional practice (see Table 2), we found that the quality of beginning teachers' classroom climate ($M = 3.21$, $SD = 0.58$) and clarity of instruction ($M = 3.14$, $SD = 0.58$) were perceived better by students compared to that of their classroom management ($M = 2.95$, $SD = 0.65$), activating learning ($M = 2.92$, $SD = 0.64$), teaching learning strategies ($M = 2.90$, $SD = 0.63$), and adaptation ($M = 2.88$, $SD = 0.57$) respectively. According to the qualification metric of teaching quality proposed by Van de Grift (2007; Maulana et al., 2013),⁷ the quality of beginning teachers' classroom climate and clarity of instruction can be interpreted as good, while the remaining domains of beginning teachers' teaching skills can be interpreted as sufficient.

5.2. Differences in teaching quality

We found that about 4%–10% of the variance in domains of

teaching quality are attributed to the school level (see Table 3). This means that there are relatively small between school differences with regard to teaching quality of beginning teachers. Furthermore, we discovered that about 42%–48% of the variance in domains of teaching quality is attributed to the teacher level. This suggests that there are large between individual beginning teachers variability in teaching quality. Finally, we found that about 44%–49% of the variance is at the occasion level.⁸ This means that there is a reasonably large variability in student perceptions of beginning teachers' teaching quality across three years.

5.3. Changes in teaching quality over time

Results of multilevel growth curve modelling show that student perceptions of all domains of teaching quality change over time (see Table 4 – 9, Model 1). In general, we found that beginning teachers' teaching quality increases during the first three years of professional practice (see Fig. 1). The change in beginning teachers' learning climate, clarity of instruction, and adaptation is best represented by a linear trend, showing that all these domains increase linearly over time ($\beta_s = 0.084, 0.091, 0.086$, $ps < 0.001$ respectively). For beginning teachers' classroom management, activating learning, and teaching strategy, the general change is also increasing over time ($\beta_s = 0.307, 0.262, 0.257$, $ps < 0.01$, respectively), but the increase is decelerated over time ($\beta_s = -0.047, -0.041, -0.040$, $ps < 0.10$ respectively). Additional analyses using a paired-samples t-test show that there is a statistically significant increase in all domains of teaching quality scores from year 1 to year 2, and from year 2 to year 3 ($ps < 0.05$). As can be seen from Fig. 1, the increase in all domains of teaching quality between year 1 and year 2 is steeper compared to between year 2 and year 3.

5.4. Between teacher differences in change over time

We found that there are between individual beginning teacher differences with regard to linear changes in teaching quality over time (see Tables 4–9, Model 2). Results show that the 95% interval contains negative as well as positive linear time effects (except for classroom

⁵ In the full model, all possible interaction effects between time and control variables as well as between variables were tested. For the reason of parsimony, only significant results were reported in this study.

⁶ In some instances, the $p < .10$ was retained to increase statistical power given the relatively moderate number of teachers in the data.

⁷ Qualification metric of teaching quality: 1.00–2.00 = insufficient, 2.01–3.00 = sufficient, 3.01–4.00 = good (Van de Grift, 2007; Maulana et al., 2013).

⁸ Part of this variance is measurement error.

Table 4
Multilevel models of Learning climate.

	Model 0		Model 1		Model 2		Model 3	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
<i>Fixed effect</i>								
Intercept	3.195***	0.020	3.039***	0.030	3.033***	0.036	2.799***	0.126
Time			0.084***	0.013	0.091***	0.014	0.165**	0.051
Time ²								
Teacher gender							–0.110	0.077
Teaching subject							0.076	0.087
Grade level							0.062	0.079
Class size							0.016	0.082
Condition							–0.107	0.077
Certification							0.405***	0.088
Time × Teacher gender							0.009	0.030
Time × Teaching subject							–0.051	0.035
Time × Grade level							–0.015	0.031
Time × Class size							0.013	0.032
Time × Condition							0.023	0.030
Time × Certification							–0.106**	0.035
<i>Random effects</i>								
Level 3 variance (school)								
Intercept	0.005	0.005	0.004	0.004	0.003	0.003	0.002	0.003
Level 2 variance (Teacher)								
Intercept	0.055	0.008	0.050	0.008	0.189	0.031	0.132	0.029
Intercept × Time					–0.053	0.012	–0.034	0.011
Time					0.016	0.005	0.010	0.005
Level 1 variance (Occasion)								
Residual	0.055	0.004	0.051	0.004	0.041	0.005	0.042	0.005
$\Delta -2\text{Loglikelihood}$		42.331***			48.832***		76.847***	

Note. + $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$.

Fig. 1. Changes in beginning teachers' teaching quality across three years (raw scores).

management only positive effects of time is visible). Recalculating the interval limits for a period of three years, the 95% interval ranges for learning climate between -0.486 and 1.032 , for classroom management between 0.105 and 1.529 , for clear instruction between -0.555 and 1.143 , for activating learning between -0.120 and 1.446 , for adaptation between -0.531 and 1.077 , and for teaching learning strategies between -0.030 and 1.338 . For all domains of teaching quality, the covariance between intercept (i.e., initial score of teaching quality) and the slope (i.e., change over time) is negative. This means that that beginning teachers who have lower initial scores in teaching

quality show a steeper increase over time than their colleagues who have higher initial scores. This pattern can also be seen in the predicted values of individual beginning teacher differences in change over time (see Fig. 2).

5.5. Contextual and teacher characteristics

We investigated the role of contextual characteristics including grade level, class size, and condition (treatment vs. control groups) in explaining differences and changes in beginning teachers'

Table 5
Multilevel models of Classroom management.

	Model 0		Model 1		Model 2		Model 3	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
<i>Fixed effect</i>								
Intercept	2.948***	0.027	2.568***	0.088	2.592***	0.087	2.477***	0.168
Time			0.307**	0.099	0.272**	0.092	0.223 ⁺	0.119
Time ²			−0.047 ⁺	0.025	−0.036	0.023	−0.032	0.027
Teacher gender							−0.102	0.089
Teaching subject							0.062	0.100
Grade level							−0.096	0.090
Class size							0.039	0.094
Condition							−0.206*	0.091
Certification							0.430***	0.101
Time × Teacher gender							0.022	0.034
Time × Teaching subject							−0.016	0.038
Time × Grade level							0.048	0.034
Time × Class size							0.007	0.035
Time × Condition							0.086**	0.033
Time × Certification							−0.086*	0.039
<i>Random effects</i>								
Level 3 variance (school)								
Intercept	0.017	0.009	0.015	0.008	0.009	0.006	0.008	0.007
Level 2 variance (Teacher)								
Intercept	0.078	0.012	0.073	0.011	0.236	0.042	0.150	0.040
Intercept × Time					−0.056	0.015	−0.029	0.015
Time					0.014	0.006	0.005	0.006
Level 1 variance (Occasion)								
Residual	0.080	0.006	0.070	0.006	0.060	0.007	0.064	0.008
Δ −2Loglikelihood		67.686***			37.989***		136.031***	

Note. ⁺ $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$.

Table 6
Multilevel models of Clear instruction.

	Model 0		Model 1		Model 2		Model 3	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
<i>Fixed effect</i>								
Intercept	3.130***	0.024	2.959***	0.035	2.952***	0.041	2.745***	0.144
Time			0.091***	0.014	0.098***	0.015	0.145**	0.057
Time ²								
Teacher gender							−0.088	0.088
Teaching subject							0.070	0.100
Grade level							0.057	0.089
Class size							0.065	0.093
Condition							−0.167 ⁺	0.087
Certification							0.361***	0.100
Time × Teacher gender							−0.003	0.034
Time × Teaching subject							−0.055	0.039
Time × Grade level							−0.006	0.034
Time × Class size							0.003	0.036
Time × Condition							0.039	0.034
Time × Certification							−0.068 ⁺	0.039
<i>Random effects</i>								
Level 3 variance (school)								
Intercept	0.011	0.007	0.008	0.006	0.006	0.004	0.000	0.000
Level 2 variance (Teacher)								
Intercept	0.072	0.011	0.066	0.010	0.237	0.038	0.187	0.037
Intercept × Time					−0.065	0.014	−0.048	0.014
Time					0.020	0.006	0.014	0.006
Level 1 variance (Occasion)								
Residual	0.066	0.005	0.062	0.005	0.048	0.006	0.049	0.006
Δ −2Loglikelihood		40.796***			43.681***		119.825***	

Note. ⁺ $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$.

teaching quality (see Tables 4–9, Model 3, Fig. 3). We found no significant effects of grade level and class size ($p > .10$). However, being in induction or non-induction groups can explain differences and changes in all domains of teaching quality except learning climate. When inspecting Fig. 3, the following patterns appear: the quality of beginning teachers' classroom management, activating learning, and teaching learning strategies in non-induction group is

better compared to induction group during the first two years of teaching practice ($p < .05$). The gap in teaching quality between the two groups became smaller in the second year. Remarkably, the opposite pattern is visible in the third year: the quality of classroom management and activating learning is better in induction group compared to non-induction group, while that of teaching learning strategies remains at about the same level in both groups.

Table 7
Multilevel models of Activating learning.

	Model 0		Model 1		Model 2		Model 3	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
<i>Fixed effect</i>								
Intercept	2.922***	0.027	2.601***	0.090	2.631***	0.087	2.500***	0.173
Time			0.262**	0.100	0.221*	0.093	0.219 ⁺	0.122
Time ²			−0.041 ⁺	0.025	−0.029	0.023	−0.027	0.027
Teacher gender							−0.107	0.092
Teaching subject							0.079	0.104
Grade level							−0.078	0.094
Class size							−0.047	0.098
Condition							−0.171 ⁺	0.094
Certification							0.460***	0.105
Time × Teacher gender							−0.009	0.035
Time × Teaching subject							−0.040	0.040
Time × Grade level							0.027	0.036
Time × Class size							0.033	0.037
Time × Condition							0.075*	0.035
Time × Certification							−0.095*	0.041
<i>Random effects</i>								
Level 3 variance (school)								
Intercept	0.016	0.009	0.013	0.008	0.007	0.006	0.006	0.006
Level 2 variance (Teacher)								
Intercept	0.084	0.012	0.078	0.012	0.258	0.044	0.175	0.042
Intercept × Time					−0.064	0.016	−0.039	0.016
Time					0.017	0.007	0.009	0.007
Level 1 variance (Occasion)								
Residual	0.078	0.006	0.072	0.006	0.060	0.007	0.065	0.008
Δ −2Loglikelihood		45.593***			38.682***		145.388***	

Note. ⁺ $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$.

Table 8
Multilevel models of Adaptation.

	Model 0		Model 1		Model 2		Model 3	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
<i>Fixed effect</i>								
Intercept	2.878***	0.023	2.718***	0.034	2.712***	0.040	2.567***	0.141
Time			0.086***	0.014	0.091***	0.015	0.116*	0.057
Time ²								
Teacher gender							−0.049	0.086
Teaching subject							0.086	0.097
Grade level							0.048	0.088
Class size							−0.051	0.091
Condition							−0.159 ⁺	0.087
Certification							0.339***	0.098
Time × Teacher gender							−0.015	0.034
Time × Teaching subject							−0.052	0.039
Time × Grade level							−0.006	0.035
Time × Class size							0.024	0.036
Time × Condition							0.048	0.034
Time × Certification							−0.058	0.040
<i>Random effects</i>								
Level 3 variance (school)								
Intercept	0.011	0.006	0.009	0.006	0.008	0.005	0.004	0.005
Level 2 variance (Teacher)								
Intercept	0.059	0.009	0.055	0.009	0.207	0.036	0.162	0.036
Intercept × Time					−0.057	0.014	−0.043	0.014
Time					0.018	0.006	0.013	0.006
Level 1 variance (Occasion)								
Residual	0.067	0.005	0.064	0.005	0.051	0.006	0.053	0.007
Δ −2Loglikelihood			35.993***		36.390***		109.081***	

Note. ⁺ $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$.

Importantly, results show that beginning teachers in the induction group display steeper growth in classroom management, activating learning, and teaching learning strategies compared to their colleagues in non-induction group across three years ($p < .05$).

To investigate the role of teacher characteristics, we included teacher certification, teaching subject, and teacher gender (see Tables 4–9, Model 3). With regard to teacher certification, we found

that certified teachers show better teaching qualities in all domains compared to uncertified teachers ($p < .05$, see Fig. 4). Although the growth in learning climate, classroom management, clarity of instruction, and activating learning is steeper for uncertified teachers, they never reach the level of certified teachers' teaching quality across three years of professional practice. Nevertheless, results also show that teaching quality of uncertified teachers increases

Table 9
Multilevel models of Teaching strategies.

	Model 0		Model 1		Model 2		Model 3	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
<i>Fixed effect</i>								
Intercept	2.901***	0.024	2.587***	0.082	2.616***	0.080	2.530***	0.154
Time			0.257**	0.092	0.218**	0.085	0.205 ⁺	0.111
Time ²			−0.040 ⁺	0.023	−0.029	0.021	−0.022	0.025
Teacher gender							−0.081	0.082
Teaching subject							0.084	0.092
Grade level							−0.025	0.083
Class size							−0.022	0.087
Condition							−0.189*	0.083
Certification							0.340***	0.093
Time × Teacher gender							−0.009	0.032
Time × Teaching subject							−0.058	0.037
Time × Grade level							0.012	0.032
Time × Class size							0.023	0.034
Time × Condition							0.066*	0.032
Time × Certification							−0.058	0.037
<i>Random effects</i>								
Level 3 variance (school)								
Intercept	0.012	0.006	0.010	0.006	0.007	0.005	0.004	0.005
Level 2 variance (Teacher)								
Intercept	0.058	0.009	0.054	0.009	0.182	0.034	0.130	0.034
Intercept × Time					−0.046	0.013	−0.029	0.013
Time					0.013	0.006	0.007	0.006
Level 1 variance (Occasion)								
Residual	0.067	0.005	0.061	0.005	0.051	0.006	0.054	0.007
Δ −2Loglikelihood			51.864***		32.011***		111.729***	

Note. ⁺ $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$.

Fig. 2. Estimated individual beginning teacher trajectories of teaching quality dimensions across three years. Horizontal axis refers to time (in years) and vertical axis is the scale score (1–4). Predicted values are from MLwiN 2.23 (Rasbach et al., 2009). Random slope parameter blocks for linear time were significant at the teacher level ($p < .01$).

over time. Results also show that the quality of certified teachers' classroom management and activating learning declines during the second and the third year of teaching practice.

Furthermore, we found that teaching subject has no significant

effects on teaching quality. Inspection of additional models without the interaction effect between time and teacher gender (not included in the Table), we found that gender could significantly explain differences in beginning teachers' teaching quality. In

(a) Classroom management

(b) Activating learning

(c) Teaching learning strategies

Fig. 3. Estimated differences in teaching quality between induction and non-induction groups based on the best-fitted multilevel growth curve models. Horizontal axis refers to time (in years) and vertical axis is the scale score (1–4).

(a) Learning climate

(b) Classroom management

(c) Clarity of instruction

(d) Activating learning

(e) Adaptation

(f) Teaching learning strategies

Fig. 4. Estimated differences in teaching quality between certified and uncertified teachers based on the best-fitted multilevel growth curve models. Horizontal axis refers to time (in years) and vertical axis is the scale score (1–4).

general, results show that the quality of teaching is better for male teachers compared to female teachers in all domains ($p < .05$), except in classroom management ($p > .10$).

6. Conclusions and discussion

The present study aims to examine variability and change in beginning teachers' teacher quality as perceived by their students during the first three years of professional practice. Additionally, the study explores the role of several personal and contextual factors in explaining levels and changes in teaching quality. The findings of the current study add to the knowledge base on the dynamics of teaching quality and on factors influencing this dynamic from the perspective of students.

6.1. Findings

6.1.1. What is the general perception of students concerning their beginning teachers' teaching quality during the first three years of professional practice?

We found that, on average, the quality of beginning teachers' classroom climate and clarity of instruction was good during the first years of teaching, while the remaining domains of beginning teachers' teaching skills were sufficient. This finding is in line with the observation research of Maulana et al. (2015). Hence, findings from student surveys and classroom observations seem to converge, suggesting that both students as well as observers detect the quality of beginning teachers' teaching in a similar way. Inexperienced teachers are better in less difficult teaching behaviour (i.e., learning climates and clear instruction), and they are still low in performing more difficult teaching behaviour (i.e., adaptation and teaching strategies). Nevertheless, another observation research of Van de Grift (2007) and Maulana et al. (2012) using samples of experienced teachers shows that teachers' adaptation to students' learning needs and teaching learning strategies remain weaker than strong, although the level is better compared to inexperienced teachers. Therefore, there seems to be a common pattern that, regardless of teaching experience, adapting teaching to meet students' learning needs and teaching learning strategies remain a complex teaching domain that needs to be improved by all teachers.

6.1.2. What are differences in perceived beginning teachers' teaching quality, between schools, between teachers, and across the first three years of professional practice?

We found a relatively small proportion of variance in teaching quality at the school level. This suggests that in terms of teaching quality, Dutch secondary schools are rather homogeneous when it comes to beginning teachers. As far as teachers' teaching quality is concerned, parents in the Netherlands may choose schools for their children wherever they like. This can be contrasted with other countries such as the United States and Indonesia where teachers' teaching quality can vary considerably across schools (Maulana et al., 2012; Jerald & Ingersoll, 2002). We also found differences between beginning teachers in the level of their teaching quality. This means that the general quality of beginning teachers' varied, ranging from the lowest to the highest quality. The fact that teaching quality varies between teachers suggests that some beginning teachers are able to maintain and improve their teaching skills after the completion of their training in teacher education, while others are still struggling. Differences within individual beginning teachers in teaching quality over time are also evident. This means that individual teacher trajectories in teaching quality varied as well. This finding supports past studies that teaching quality is not a stable characteristic. It is dynamic that changes over

time (Maulana et al., 2013; Opdenakker et al., 2012) influenced by changing circumstances.

6.1.3. Does student perception of beginning teachers' teaching quality change during the first three years of professional practice?

We found that all domains of beginning teachers' teaching quality increased across the first three years of teaching. This finding is in line with the research of Henry, Bastian, and Fortner (2011) who found that average beginning American teachers' effectiveness increases during the first three years of teaching. Over time, the developmental level of activating learning, adaptation, and teaching learning strategy never reached that of learning climate, clarity of instruction, and classroom management. This finding supports the observation research of Maulana et al. (2015) regarding the increasing difficulty level of teaching domains, studied from the longitudinal perspective. They showed that the six teaching domains follow a complexity ranking, ranging from less complex (learning climate, clarity of instruction, classroom management) to more complex teaching behaviors (activating learning, teaching learning strategy, adaptation). In all domains of teaching quality, we found a steeper increase during the first two years of teaching, then the increase was decelerated slightly towards the third year (this is particularly visible for classroom management, activating learning, and teaching strategy). The steeper (earlier) increase in classroom management was also found in the observed teaching quality of English beginning teachers (Malmberg et al., 2010). Additionally, the developmental patterns of teaching quality found in the current study are in line with the literature on supported beginning and in-service teachers, which suggested a steeper initial increase in teaching quality and a tendency to level-off towards the end-of-time-span (Hebert & Worthy, 2001; Woolfolk-Hoy & Spero, 2005). Unlike the research of Veenman (1984) indicating a "reality shock" during the entry to profession characterized by a dip in teaching quality, our research suggests that beginning teachers were progressively able to overcome their concerns to some extent during the beginning phase of their career.

Why do the six domains of teaching quality increase over time? One line of reasoning could be that the first year of teaching is usually filled with optimism and commitment, although at the same time this period is often experienced as stressful (Hebert & Worthy, 2001; Woolfolk-Hoy & Spero, 2005). Research recognizes that teachers entering the profession became less democratic and more custodial over time (Hoy & Woolfolk, 1990). However, there is also evidence that teachers show increases in organized, systematic, affective and stimulating teaching behaviour (Adams, 1982). An interview study from Australia indicates that beginning teachers developed a greater capacity for reflection in their first year of teaching (Cavanagh & Prescott, 2010), which is useful for their learning to develop better teaching quality subsequently, as reflected in the current study. Activating learning, teaching-learning strategy and adaptation of teaching remain the most difficult domains of teaching for beginning teachers. However, these three teaching quality domains increase over time, too.

6.1.4. Are there differences between individual beginning teachers with regard to change across 3 years as perceived by students?

Our multilevel model indicates that beginning teachers' teaching quality differed in their development over time. This intrapersonal variability in the development of teaching quality can be a result of teacher adaptation to or coping with situational demands (Frenzel & Götz, 2007; Gibson & Dembo, 1984). It can also be related to intrapersonal aspects of teachers such as their motivation and morale, adaptability for changing in working conditions, age and developmental phase, self-efficacy, and satisfaction through student progress and achievement (Day, 2004; Fessler, 1995; Fullan & Hargreaves,

1992; Rosenholtz, 1989; Troman & Woods, 2001; Tschannen-Moran et al., 1998) The covariance structure suggests a mastery effect over time. In all domains of teaching quality, teachers who started off at a low level increased more over time than teachers who started off at a higher level. This finding is in line with observed classroom management of English beginning teachers (Malmberg et al., 2010). As Bloom (1971) asserted, under supportive conditions, individuals become more similar in their level of achievement and in their learning gain. Because beginning teachers in the current study show this tendency, it seems to be that the school context where the beginning teachers worked during their first three years of teaching was appropriately supportive (i.e., supportive mentors). With mastery learning, beginning teachers who fall into faster and slower learners could increase their teaching quality. Slower learners could benefit from a more significant increase. Mastery learning is one way slow learners can be helped to increase their performance (Guskey & Gates, 1986).

6.1.5. How do contextual features (class size, grade level, induction program) and beginning teacher characteristics (teaching subject, gender, certification) explain differences and changes in beginning teachers' perceived teaching quality?

We found that grade level and class size could not explain differences in teaching quality. This indicates that teaching quality of beginning teachers was similar regardless of in which grade they taught and how big the number of students in the class was. This finding seems to support the minimal effect of class size on teaching quality (Ehrenberg, Brewer, Gamoran, & Willms, 2001). Regardless of this finding, effects of class size on teaching quality should not be underestimated. There is evidence from the English context that smaller classes benefit all students because of individual attention from the teacher. In secondary education particularly, low-attaining students benefit more from small classes. Students in large classes, on the other hand, drift off task because of too much instruction from the teacher to the whole class instead of individual attention, which has a negative consequence for low-attaining students (Blatchford, Bassett, & Brown, 2011). However, reducing class size does not automatically mean that teaching quality will be improved. Research from the American context indicated that in smaller classes, too, teachers need to master effective strategies to create meaningful learning opportunities for students (Graue, Hatch, Rao, & Oen, 2007).

Teacher induction could explain differences and changes in teaching quality over time. Beginning teachers in the induction group showed a much faster growth in the quality of classroom management, activating learning, and teaching learning strategies compared to their colleagues in the non-induction group over time. In the third year, beginning teachers in the induction group outperformed their colleagues in the non-induction group in terms of the quality of classroom management and activating learning (and teaching learning strategies to a minor extent). Our findings confirm the beneficial effect of teacher induction program for beginning teachers. As in the current study, the induction program was targeted at providing social and professional support for beginning teachers (Beijaard, Buitink, & Kessels, 2010). Social and professional supports, packed in the teacher induction program, are important for the development of teachers in schools (Maulana et al., 2015). Our finding is in line with past research revealing that the induction program was related to increased reflection, improved collegiality, openness and communication, greater teacher autonomy, self-growth, and personal efficacy among English beginning teachers (Smetthem & Adey, 2005). The developmental trend of beginning teachers in the non-induction group, on

the other hand, is more in line with Henry et al. (2011) indicating that teaching quality will tend to flatten after three years of professional practice. The induction program seems to be powerful in mitigating this plateauing trend during the early career.

Regarding teacher characteristics, we found that teaching subject had no significant effect on beginning teachers' teaching quality. However, we found that certification status and teacher gender could explain differences and changes in teaching quality. Certification status matters when it comes to teaching quality, as our findings revealed that certified teachers displayed higher quality of teaching (in all six domains) than uncertified teachers. This finding is in agreement with past studies in the United States showing that uncertified teachers did not perform as well as certified teachers, even after controlling for teaching experience (Fetler, 1997). Students taught by certified teachers had higher (math) achievement (Fuller & Alexander, 2004), while students taught by uncertified teachers had the lowest (math) achievement (Darling-Hammond, 2000). The fact that certification status matters for teacher development and student outcomes suggests that teacher preparation forms an essential requirement before teachers enter the profession. This is signified by a review study of Greenwald et al. (1996) showing that teachers with pedagogical training performed better than those who entered teaching without such training. Additionally, we found that in general male teachers performed better teaching quality than female teachers. This finding is in line with past studies with various teaching experience in Dutch and Flemish contexts showing that compared to female teachers, male teachers displayed better leadership and cooperativeness with students (Van Petegem, Creemers, Rosseel, & Aelterman, 2005; Opdenakker et al., 2012) and showed better classroom management (Opdenakker & van Damme, 2007).

6.2. Implications for policy and practice

The present study suggests that it takes up to three years after entering the teaching profession for beginning teachers to reach a strong level of teaching quality in all domains. During the first two years of teaching, particularly, the quality of classroom management, activating learning, adaptation to students' learning needs, and teaching learning strategy requires a significant improvement. These findings imply the extent to which teacher education plays a role in preparing pedagogical-didactical training of pre-service teachers and the role of school support system (i.e., mentors) during the first three years of professional practice. Generally, (Dutch) teacher education graduates tend to have a strong level of less complex teaching domains (learning climates and clarity of instruction). Only after three years of teaching, the more complex domains of teaching quality were improved to a strong level. Perhaps, this is not surprising given that professional standards previously in force in the Netherlands put little emphasis on complex aspects of teaching such as adaptation to students' learning needs and teaching learning strategy. If educational outcomes are to be enhanced, teacher education and school support system need to put emphasis on promoting higher levels of teaching beyond stimulating learning climates and displaying instructional clarity.

The challenge for teacher education and school mentors is how to help beginning teachers reach a strong level of teaching quality earlier in the career. For teacher education, it is important to set minimum goals and determine reasonable passing criteria for all domains of teaching quality. Results of the current study can be used as a general guideline for setting such criteria (see Table 1). For example, levels of teaching quality at time 1 can be used as a norm

for pre-service teachers to officially enter teaching profession. A more ambitious teacher education can set higher standards by, for instance, using levels of teaching quality at time 3 (as shown in the current study) as the norm. As our study indicated, teacher induction programs are one effective way of boosting teachers' teaching quality during the early career. The development of beginning teachers' teaching quality can be supported in the form of extensive mentoring that last through the first years of teaching (Scherer, 2012; Shagrir, 2012). In particular, mentoring that follows constructivist rather than transmissive principles of learning would foster the growth of teacher efficacy, teaching enthusiasm, and job satisfaction and reduces emotional exhaustion (Richter et al., 2013). Because research has consistently shown that beginning teachers perceived their mentors as one of their most important source of support during the first year of teaching (Lindgren, 2005; Marable & Raimondi, 2007), this implies that efforts to train teacher educators and school mentors should be a priority in the (inter)national agenda of educational improvement.

Finally, the current study is in line with past studies documenting changes in teaching quality over time (i.e., Maulana et al., 2012; Maulana et al., 2013; Malmberg et al., 2010; Opdenakker et al., 2012), which suggests that teaching quality is a dynamic and malleable factor. However, other studies discovered that teaching quality is rather stable across three and six lessons (Kunter & Baumert, 2006; Seidel et al., 2002). Comparing the mentioned studies together, there seems to be a general trend that changes in teaching quality can be visible within a certain period (i.e., beyond the six lesson period). This suggests that behavioural changes take time and that efforts for tracking changes in teaching quality should be extended beyond six lesson intervals. For teacher educators, school coaches, and school mentors, this implies that observing beginning teachers for evaluation purposes should be scheduled with ample intervals. If the development of mastering teaching skills is of importance, short-term development (i.e., within-year changes) can be observed by gathering data in the beginning and the end of the school year. For longer-term development (i.e., between-year changes), the longitudinal design such as described in the present study can be used.

6.3. Limitations and suggestions for future research

Although our study involves a large national sample within a longitudinal design, the study should be considered in light of a few limitations. First, our measurement was limited to the first three years of teaching. Although there is a tendency that the development of teaching quality was less steep after the first year of teaching, the moments when teaching quality would flatten remains inconclusive. Information about the flattening moments of teaching quality during the first years of teaching is important because it may suggest that an additional and different support for improving teaching quality may be necessary. When feasible, future research should extend the measurement of the development of teaching quality beyond the first three years of teaching. Second, interpretations of teaching quality of the current study was limited to perceptions of students. Although students' perceptions are a valid source of information regarding their teachers' teaching quality (Brok et al., 2004; Fraser, 2002), future research should try to collect additional data from expert observers to validate the current findings. Thirdly, this study does not cover a broader conceptualization of teaching quality. Future research should include broader measures of teaching quality to measure effects in a wider scope.

6.4. Conclusion

In conclusion, the current study suggests that beginning teachers' teaching quality gets better during the first three years of professional practice. Nevertheless, the general level of their teaching quality is still relatively weak. Furthermore, the growth in teaching quality is steeper during the first two years of teaching. Knowing that changes in teaching quality during the beginning of professional practice are visible, in combination with the finding that beginning teachers' teaching quality is just sufficient generally, better teacher preparation during pre-service teacher education and more supportive guidance from school mentors in schools are called for. Particular attention needs to be given to supporting the acceleration of teaching quality after the second year of professional teaching onwards. In line with findings of Helms-Lorenz, Van de Grift, and Maulana (2015), the acceleration of growth in teaching quality can be done by paying ample attention to teaching skills at the start, classroom observation, coaching, and reducing workload.

Acknowledgements

Our gratitude goes to Anna Verkade for her great assistance during the implementation of the national and longitudinal study. Our special thanks also go to Jaap Buitink, Chrisbert van Mourik, Geke Schuurman, and Wiepkje t'Hart for their great commitment and dedication to the national induction program. This study was supported financially by Dutch ministry of Education via Actieprogramma Onderwijs Bewijs [Action Program Education Evidence, grant number: ODB08110].

Appendix 1

Student questionnaire measuring teachers' teaching quality.

Domain	Item content
	My teacher ...
Learning climate	... treats me with respect
	... answers my question
	... ensures that I treat others with respect
Classroom management	... applies clear rules
	... ensures that I pay attention
	... ensures that I use my time effectively
Clear instruction	... prepares his/her lesson well
	... makes clear what I need to study for a test
	... uses clear examples
Activating learning	... explains the purpose of the lesson
	... explains everything clearly to me
	... involves me in the lesson
Adaptive teaching	... ensures that I do my best
	... motivates me
	... helps me if I do not know, or am unable to do something
	... helps me if I do not understand or am unable to do something
	... pays attention to me
	... connects to what I know or am capable of
Teaching learning strategies	... checks whether I have understood the subject matter
	... makes me feel self-confident with difficult tasks
	... knows what I have difficulty with
	... explains how I need to do things
	... stimulates my thinking
	... explains how I should study something

Appendix 2

Guidelines to optimize working conditions; ensuring instructive work and reducing work pressure.

Goals	Measure that leads to less stress	Examples of activities
Optimizing organizational preconditions and optimizing the working situation Tailor-made measures: fine-tuning of work pressure, challenges and capacity	No extra tasks	<ul style="list-style-type: none"> ■ No tutorship (build up from 1 to 3); also made-to-measure; starting point as basis ■ No extra tasks ■ Reduced number of lessons ■ Discuss perception of work pressure with coach ■ Learn time-management ■ Organize work efficiently, work goal-oriented ■ Learn to be a chairman ■ Keep guard over time, also after school ■ take care of tailor-made assignments: ■ provide the opportunity to reflect on workload ■ give possibility to change classes ■ make and keep tasks suitable ■ teach parallel classes preferably with experienced colleague ■ preferably no scheduled lessons at the end of the day ■ schedule weekly hour for support ■ secure confidentiality between BT and support provider ■ attune schedules of BTs and the subject coach ■ Communicate well with schedules makers ■ Reduce excessive commuting between buildings ■ Provide BTs with own classrooms
Give no assignments that can't be accomplished: not too easy and not too difficult assignments.	Attune work to BT's capacity	
	Advantageous schedules	<ul style="list-style-type: none"> ■ Provide a quiet and good workplace ■ Provide continual "fitting" challenges ■ Provide accessible resources ■ Placement in dynamic, helpful teams ■ Trying to have the supporting teacher and the coach in the same team ■ Organize weekly (peer) reflection sessions
Acknowledgement of the fact that development takes time Instructive workplace	Develop induction arrangements for the duration of three years and see to a gradual increase of workload up to the normal end level. Placement in particular classes and teams.	

Appendix 3

Guidelines to support the development of the beginning teacher.

Domains	Developmental goals	Guidelines	Content of activities to accomplish goals (what?)	Activity (how?)
School organization and context	To gain specific background knowledge of students, the curriculum, the school and the social context.	Provide opportunity for informal, unintentional learning Take school policy into account when developing an induction arrangement Providing a coach of the same school and same subject field (matching teaching conceptions between mentee and mentor increases effectiveness) Provide opportunities for formal learning BT who had an appointment during their internship should be treated in the same way as BT who had their internship elsewhere.	<ul style="list-style-type: none"> ■ Getting to know the content of student files (study progress and personal information) and learning how to use this information correctly ■ Knowing what the roles of teachers is ■ Getting to know the social context of the average student ■ Knowing the mission, the identity and the uniqueness of the school ■ Getting to know the structure, organization and culture of all sections/departments within school ■ Being aware of pupil transition problems inside the school. ■ Being informed about current social discussions concerning education 	<ul style="list-style-type: none"> ■ Appoint a (senior) teacher to be responsible for BTs; , for instance a school educator or a team leader ■ Activities directed to introduce the use student files ■ Having informal conversations with students ■ Have BTs participate with school trips, not as organizers ■ Supply information about school structure, organisation and culture. ■ Provide opportunities to attend different kinds of meetings ■ Be transparent about decisions made by the school concerning the organization of learning processes. ■ Let BT's visit other teams, locations, levels, departments in- and outside of the school ■ Make social developments accessible and discussible for instance by using

(continued)

Domains	Developmental goals	Guidelines	Content of activities to accomplish goals (what?)	Activity (how?)
Effective teaching behaviour	Cultivation of consciousness of the fact that you are a part of a team and an institution and knowing your responsibilities	Provide the opportunity for the BT to be a full member of the institution (in rights and duties)	<ul style="list-style-type: none"> ■ Getting to know the values and norms of the organization and its members ■ Getting to know the explicit and implicit rules of the organization ■ Learning to meet effectively 	journal articles, newspaper articles, subscription to journals, either on paper or on internet <ul style="list-style-type: none"> ■ See to it that social developments are discussed during meetings ■ Coach collegiality ■ Provide transparency ■ quality assurance system ■ Learning effective time-management skills ■ Attending meetings and workshops ■ Learning effective meeting skills ■ Discuss how to behave as a colleague during meetings and in the teams ■ Critical reading of learning theories from scientific publications and literature ■ Using think-out-loud-protocols of students and analysing these ■ Attending workshops about learning and teaching
	Focus on learning of the students as a starting point of teaching behaviour	Stimulate lifelong learning	<ul style="list-style-type: none"> ■ Knowledge about theories of learning processes and integration thereof into the subject ■ Continuous development of the notion of “how learning works” ■ Noticing (assessing, observing) if and how students learn 	<ul style="list-style-type: none"> ■ Consulting scientific literature ■ Attending courses/extra training/congresses ■ Consulting different teaching methods ■ Attending thematic workshops or workshops with experts ■ coaching on the job ■ study literature and other relevant information ■ develop evidence-based teaching ■ specific observation and processing of feedback ■ Watching and discussing video recordings of oneself and others ■ Preparing lessons together: in pairs and/or in teams ■ Asking feedback on a plan for a lesson ■ Workshop about learning to develop strategies for teaching ■ Observing the behaviour of an expert ■ Giving space for experimentation through mutual observations in lessons
	Deepening and extending subject content knowledge	Lifelong learning Building on a knowledge base that transcends the own subject matter	<ul style="list-style-type: none"> ● Keeping track of latest knowledge of the subject and latest developments 	
	Applying, extending and refining beginning behaviour repertoire	Provide opportunities for reflection on own behaviour Developing a notion of how own classroom behaviour and how this effects pupils	<ul style="list-style-type: none"> ■ Being informed about new developments in didactics ■ Noticing differences between students (backgrounds, behaviour, development) ■ Learning to treat (groups of) individuals differently in order to enhance overall learning. ■ Learning which differences call for didactic differentiation ■ Improve effectiveness of own teaching ■ Intensifying interaction with and between students 	
	Knowledge about student counselling and being able to function as a tutor	Developing knowledge of counselling and exercising the skills	<ul style="list-style-type: none"> ■ Introduction to student counselling in general and in this school specific ■ Develop communication skills ■ Getting to know the limits of tutorship 	
	The development of an effective learning climate in the classroom, for instance creating a safe pedagogical climate Being aware of your role	Negotiating with students about the learning climate Consulting with colleagues, reflection Observation and feedback Translate theory to practice	<ul style="list-style-type: none"> ■ Being informed about the rules of the school, protocols and being able to play with it ■ Knowing theory about motivation, group behaviour, problem behaviour, bullying etc. ■ Being informed about group dynamics ■ Setting boundaries/limits ■ Knowledge of effects of own behaviour ■ Managing individual differences ■ Being able to name values and norms in the classroom and taking care of the execution of the rules 	<ul style="list-style-type: none"> ■ Getting training to be a tutor and growing into tutorship ■ Training of conversation skills for conversations with students ■ Being Informed about specific issues as incest, domestic violence etc. ■ Introduction to school rules ■ Consulting and using literature ■ Talking with tutors ■ Attending thematic workshops ■ Having conversations with students about values and norms. Exchanging thoughts with colleagues about this. Making consequences explicit.
	Knowledge of the individual differences between students and their diversity Skills to manage these differences Special attention for and knowledge of the problems of students with special needs	Acknowledgement of the complexity of a heterogeneous group	<ul style="list-style-type: none"> ■ Understanding problem behaviour like ADHD, PDDNOS, dyslexia, dyscalculia ■ Knowledge of facilities to support special needs ■ Being informed about behaviour protocols like bullying behaviour and coping with grief 	<ul style="list-style-type: none"> ■ Making background information available, journals etc. ■ Consulting these sources ■ Attending student discussions in teams ■ Coordinator of care gives information about set-up, content and structure of care

(continued on next page)

(continued)

		Descriptive statistics evaluation										
		% Missing	Extreme values		t-test table*							
			Low	High	Teaching subject (Mean)		Class size (Mean)		Grade level (Mean)		Certification (Mean)	
					Present	Missing	Present	Missing	Present	Missing	Present	Missing
Grade level ^d	19.6	—	—	—	—	—	—	—	—	—	—	
Teacher gender ^e	0.9	—	—	—	—	—	—	—	—	—	—	
Certification ^f	9.5	—	—	—	—	—	—	—	—	—	—	
Time 2												
Dependent variables												
Learning climate	42	2	0	3.21	3.33	3.21	3.33	3.21	3.33	3.11	3.24	
Classroom management	42	3	0	2.98	3.09	2.98	3.09	2.98	3.09	3.03	3.06	
Clear instruction	42	2	0	3.15	3.30	3.15	3.30	3.15	3.30	3.11	3.18	
Activating learning	42	0	0	2.94	3.09	2.94	3.09	2.94	3.09	2.94	2.95	
Adaptation	42	0	0	2.89	3.02	2.89	3.02	2.89	3.02	2.85	2.84	
Teaching strategies	42	2	0	2.92	3.06	2.92	3.06	2.92	3.06	2.79	2.81	
Independent variables												
Teaching subject ^a	19.9	—	—	—	—	—	—	—	—	—	—	
Class size ^b	19.6	—	—	—	—	—	—	—	—	—	—	
Condition ^c	19.6	—	—	—	—	—	—	—	—	—	—	
Grade level ^d	0.9	—	—	—	—	—	—	—	—	—	—	
Teacher gender ^e	9.5	—	—	—	—	—	—	—	—	—	—	
Certification ^f	9.5	—	—	—	—	—	—	—	—	—	—	
Time 3												
Dependent variables												
Learning climate	52.7	3	0	3.31	3.35	3.31	3.35	3.21	3.33	3.21	3.18	
Classroom management	52.7	5	0	3.08	3.16	3.08	3.16	2.98	3.09	3.11	3.10	
Clear instruction	52.7	3	0	3.24	3.30	3.24	3.30	3.15	3.30	3.19	3.20	
Activating learning	52.7	2	0	3.02	3.14	3.02	3.14	2.94	3.09	3.01	3.06	
Adaptation	52.7	1	0	2.97	3.06	2.97	3.06	2.89	3.02	2.92	2.93	
Teaching strategies	52.7	2	0	3.00	3.07	3.00	3.07	2.92	3.06	2.86	2.86	
Independent variables												
Teaching subject ^a	19.9	—	—	—	—	—	—	—	—	—	—	
Class size ^b	19.6	—	—	—	—	—	—	—	—	—	—	
Condition ^c	19.6	—	—	—	—	—	—	—	—	—	—	
Grade level ^d	0.9	—	—	—	—	—	—	—	—	—	—	
Teacher gender ^e	9.5	—	—	—	—	—	—	—	—	—	—	
Certification ^f	9.5	—	—	—	—	—	—	—	—	—	—	

Note. ^a 0 = science, 1 = non-science, ^b 0 = small class, 1 = large class, ^c 0 = control group, 1 = induction group, ^d 0 = lower grade, 1 = higher grade, ^e 0 = male, 1 = female, ^f 0 = uncertified, 1 = certified. * Indicators with less than 5% missing are not reported.

References

- Adams, R. D. (1982). A look at changes in teacher perceptions and behavior across time. *Journal of Teacher Education*, 33, 40–43. <http://dx.doi.org/10.3102/00346543054002143>.
- Alderman, G. L., & Nix, M. (1997). Teachers' intervention preferences related to explanations for behavior problems, severity of the problem, and teacher experience. *Behavior Disorders*, 22, 87–95.
- Algozzine, B., Grete, J., Queen, A. J., & Cowan-Hathcock, M. (2007). Beginning teachers' perceptions of their induction program experiences. *Clearing House: A Journal of Educational Strategies, Issues and Ideas*, 80(3), 137–143. <http://dx.doi.org/10.3200/TCHS.80.3.137-143>.
- Beijaard, D., Buitink, J., & Kessels, C. (2010). Teacher induction. In P. L. Peterson, E. Baker, & B. McGaw (Eds.), *International encyclopaedia of education* (Vol. 7, pp. 563–568). Oxford: Elsevier.
- Blatchford, P., Bassett, P., & Brown, P. (2011). Examining the effect of class size on classroom engagement and teacher-pupil interaction: differences in relation to pupil prior attainment and primary vs. secondary education. *Learning and Instruction*, 21, 715–730. <http://dx.doi.org/10.1016/j.learninstruc.2011.04.001>.
- Bloom, B. S. (1971). *Individual differences in school achievement: A vanishing point?* (Phi Delta Kappan Monograph). Bloomington: Ind: Phi Delta Kappan International.
- Brok, P. den., Bergen, Stahl, R. J., & Brekelmans, M. (2004). Students' perceptions of teacher control behaviours. *Learning and Instruction*, 14, 425–443.
- Carnine, D. W., Dixon, R. C., & Silbert, J. (1998). Effective strategies for teaching Mathematics. In E. J. Kameenui, & D. W. Carnine (Eds.), *Effective teaching strategies that accommodate diverse learners*. Upper Saddle River, NJ: Prentice Hall.
- Cavanagh, M., & Prescott, P. (2010). The growth of reflective practice among three beginning secondary mathematics teachers. *The Asia-Pacific Journal of Teacher Education*, 38(2), 147–159. <http://dx.doi.org/10.1080/13598661003678968>.
- Chen, Q. (2012). The impact of ignoring a level of nesting structure in multilevel mixture model: a monte carlo study. *Sage Open*, 2, 1–10. <http://dx.doi.org/10.1177/2158244012442518>.
- Choy, D., Wong, A. F. L., Lim, K. M., & Chong, S. (2013). Beginning teachers' perceptions of their pedagogical knowledge and skills in teaching: a three year study. *Australian Journal of Teacher Education*, 38, 68–79. <http://dx.doi.org/10.14221/ajte.2013v38n5.6>.
- Cornelius-White, J. (2007). Learner-centered teacher-student relationships are effective: a meta-analysis. *Review of Educational Research*, 77, 113–143. <http://dx.doi.org/10.3102/003465430298563>.
- Creemers, B. P. M., & Kyriakides, L. (2008). *The dynamics of educational effectiveness: A contribution to policy, practice and theory in contemporary schools*. London: Routledge.
- Danielson, C. (2013). *The framework for teaching: Evaluation instrument*. Princeton: The Danielson Group.
- Darling-Hammond, L. (2000). Teacher quality and student achievement: a review of state policy evidence. *Education Policy Analysis Archives*, 8, 1–44.
- Darling-Hammond, L., Chung, R., & Frelow, F. (2002). Variation in teacher preparation: how well do different pathways prepare teachers to teach? *Journal of Teacher Education*, 53, 286–302. <http://dx.doi.org/10.1177/0022487102053004002>.
- Day, C. (1999). *Developing teachers: The challenges of lifelong learning*. London: Falmer Press.
- Day, C. (2004). *A passion for teaching*. London: Routledge Falmer.
- De Jong, R., & Westerhof, K. (2001). The quality of student ratings of teacher behaviour. *Learning Environments Research*, 4, 51–85. <http://dx.doi.org/10.1023/A:1011402608575>.
- Deci, E. L., & Ryan, R. M. (2000). The 'what' and 'why' of goal pursuits: human needs and the self-determination of behavior. *Psychological Inquiry*, 11(4), 227–268. http://dx.doi.org/10.1207/S15327965PLI1104_01.
- Ehrenberg, R. G., Brewer, D. J., Gamoran, A., & Willms, J. D. (2001). Class size and student achievement. *Psychological Science in the Public Interest*, 2(1), 1–30.
- Feiman-Nemser, S. (2001). From preparation to practice: designing a continuum to strengthen and sustain teaching. *Teachers College Record*, 103(6), 1013–1055.
- Feiman-Nemser, S., Schwill, S., Carver, C., & Yusko, B. (1999, July). *A conceptual review of literature on new teacher induction*. Paper presented at the meeting of the National Partnership for Excellence and Accountability in Teaching,

- and school leadership affect school practice and outcomes in secondary education? *British Educational Research Journal*, 33(2), 179–206. <http://dx.doi.org/10.1080/01411920701208233>.
- Pearson, P. D., & Fielding, L. (1991). Comprehension instruction. In R. Barr, M. L. Kamil, P. B. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research* (Vol. II, pp. 815–860). White Plains, NY: Longman.
- 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>.
- Pressley, M., Wood, E., Woloshyn, V. E., Martin, V., King, A., & Menke, D. (1992). Encouraging mindful use of prior knowledge. Attempting to construct explanatory answers facilitates learning. *Educational Psychologist*, 27(1), 91–109. http://dx.doi.org/10.1207/s15326985ep2701_7.
- Rasbash, J., Charlton, C., Browne, W. J., Healy, M., & Cameron, B. (2005). *MLwiN (Version 2.0) [Computer software]*. Bristol, UK: University of Bristol.
- Rasbash, J., Steel, F., Browne, W., & Goldstein, H. (2014). *A user's guide to MLwiN (Version 2.1)*. Centre for Multilevel Modelling. University of Bristol.
- Reis, S. M., McCoach, D. B., Little, C. A., Muller, L. M., & Kaniskan, R. B. (2011). The effects of differentiated instruction and enrichment pedagogy and reading achievement in five elementary schools. *American Educational Research Journal*, 48, 462–501. <http://dx.doi.org/10.3102/0002831210382891>.
- Richter, D., Kunter, M., Lüdtke, O., Klusmann, U., Anders, Y., & Baumert, J. (2013). How different mentoring approaches affect beginning teachers' development in the first years of practice. *Teaching and Teacher Education*, 36, 166–177. <http://dx.doi.org/10.1016/j.tate.2013.07.012>.
- Rivkin, S. G., Hanushek, E. A., & Kain, J. F. K. (2005). Teachers, schools, and academic achievement. *Econometrica*, 73, 417–458.
- Rosenholtz, S. (1989). *Teachers' workplace: the social organisation of school*. New York: Longman.
- Rosenshine, B. (1980). How time is spend in elementary classrooms. In C. Denham, & A. Lieberman (Eds.), *Time to learn* (pp. 107–126). Washington D.C: National Institute of Education.
- Rosenshine, B. V., & Meister, C. (1997). Cognitive strategy instruction in reading. In S. Stahl, & D. A. Hayes (Eds.), *Instructional models in Reading* (pp. 85–107). Mahwah, NJ: Erlbaum.
- Rosenshine, B. V., & Stevens, R. (1986). Teaching functions. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 376–390). New York: Macmillan.
- Scantlebury, K., Boone, W., Kahle, J. B., & Fraser, B. J. (2001). Design, validation, and use of an evaluation instrument for monitoring systemic reform. *Journal of Research in Science Teaching*, 38, 646–662. <http://dx.doi.org/10.1002/tea.1024>.
- Scheerens, J., & Bosker, R. (1997). *The foundations of educational effectiveness*. Oxford: Pergamon.
- Scherer, M. (2012). The challenges of supporting new teachers: a conversation with Linda Darling-Hammond. *Educational Leadership*, 69, 18–23.
- Seidel, T., Prenzel, M., Duit, R., Euler, M., Geiser, H., Hoffmann, L., et al. (2002). "Jetzt bitte alle nach vorne schauen!" – Lehr-Lernskripts im Physikunterricht und damit verbundene Bedingungen für individuelle Lernprozesse ["Can everybody look tot he front of the class please!" – Patterns of instruction in elementary physics classroom and its implication for students' learning]. *Unterrichtswissenschaft*, 30(1), 52–77.
- Seidel, T., & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: the role of theory and research design in disentangling meta-analysis results. *Review of Educational Research*, 77, 454–499. <http://dx.doi.org/10.3102/0034654307310317>.
- Shagrir, L. (2012). How evaluation processes affect the professional development of five teachers in higher education. *Journal of the Scholarship of Teaching and Learning*, 12(1), 23–35.
- Sijtstra, J. (1997). *Balans van het taalonderwijs aan het einde van de basisschool: Uitkomsten van de tweede taalpeiling einde basisonderwijs. [Achievements in mother tongue at the end of primary education.]*. Arnhem: CITO.
- Slavin, R. (1996). *Education for all*. Exton, PA: Swets & Zeitlinger Publishers.
- Smethem, L., & Adey, K. (2005). Some effects of statutory induction on the professional development of newly qualified teachers. A comparative study of pre- and post-induction experiences. *Journal of Education for Teaching*, 31, 187–200. <http://dx.doi.org/10.1080/02607470500169014>.
- Smith, T. W., Baker, W. K., Hattie, J. A. C., & Bond, L. (2008). A validity study of the certification system of the National Board for Professional Teaching Standards. In L. Ingvarson, & J. A. C. Hattie (Eds.), *Assessing teachers for professional certification: The first decade of the National Board for Professional Teaching Standards* (pp. 345–378). Oxford, UK: Elsevier.
- Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel Analysis: An introduction to basic and advanced multilevel modelling* (2nd ed.). London: Sage.
- Sykes, B., & Kuyper, H. (2013). School segregation and the secondary school achievements of youth in the Netherlands. *Journal of Ethnic and Migration Studies*, 39, 1699–1716. <http://dx.doi.org/10.1080/1369183X.2013.833707>.
- Teodorovic, J. (2011). Classroom and school factors related to student achievement: what works for students? *School Effectiveness and School Improvement*, 22, 215–236. <http://dx.doi.org/10.1080/09243453.2011.575650>.
- Tomlinson, C. (2003). *Fulfilling the promise of the differentiated classroom: Tools and strategies for responsive teaching*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Tomlinson, C. (2005). Grading and differentiation: paradox or good practice? *Theory into Practice*, 44, 262–269. http://dx.doi.org/10.1207/s15430421tip4403_11.
- Troman, G., & Woods, P. (2001). *Primary teachers' stress*. London: Routledge.
- Tschannen-Moran, M., Hoy, A. W., & Hoy, W. (1998). Teacher efficacy: its meaning and measure. *Review of Educational Research*, 24, 9–17. <http://dx.doi.org/10.3102/00346543068002202>.
- Van Petegem, K., Creemers, B. P. M., Rosseel, Y., & Aelterman, A. (2005). Relationships between teacher characteristics, interpersonal teacher behaviour and teacher wellbeing. *Journal of Classroom Interaction*, 4, 34–43. <http://dx.doi.org/10.1007/s11205-006-9055-5>.
- Van de Grift, W. J. C. M. (2007). Quality of teaching in four European countries: a review of the literature and application of an assessment instrument. *Educational Research*, 49, 127–152. <http://dx.doi.org/10.1080/00131880701369651>.
- Veenman, S. (1984). Perceived problems of beginning teachers. *Review of Educational Research*, 54, 143–178. <http://dx.doi.org/10.3102/00346543054002143>.
- Wang, M. C., Reynolds, M. C., & Walberg, H. J. (1995). Serving students at the margins. *Educational Leadership*, 52(4), 12–17.
- Willms, J. D., & Somer, M.-A. (2001). Family, classroom, and school effects on children's educational outcomes in Latin America. *School Effectiveness and School Improvement*, 12, 409–445. <http://dx.doi.org/10.1076/sesi.12.4.409.3445>.
- Wilson, S. M., Floden, R. E., & Ferrini-Mundy, J. (2002). Teacher preparation research: an insider's view from the outside. *Journal of Teacher Education*, 53, 190–204. <http://dx.doi.org/10.1177/0022487102053003002>.
- Woolfolk-Hoy, A. W., & Spero, R. B. (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>.
- Worthen, B. R., Sanders, J. R., & Fitzpatrick, J. L. (1997). *Program evaluation: An alternative approaches and practical guidelines*. White Plains, NY: Longman.
- Yair, G. (2000). Educational battlefields in America: the tug-of-war over students' engagement with instruction. *Sociology of Education*, 73(4), 247–269.
- Yeaman, A., Hlynka, D., Anderson, J., Damarin, S., & Muffoletto, R. (1996). Post-modern and poststructuralist theory. In D. Jonassen (Ed.), *Handbook of research for educational communications and technology* (pp. 226–252). New York: Simon and Shuster Macmillan.