

University of Groningen

## How Features of the Implementation Process Shape the Success of an Observation-Based Coaching Program

Hu, Yanjuan; van Veen, Klaas

*Published in:*  
Elementary School Journal

*DOI:*  
[10.1086/711070](https://doi.org/10.1086/711070)

**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:*  
2020

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

*Citation for published version (APA):*

Hu, Y., & van Veen, K. (2020). How Features of the Implementation Process Shape the Success of an Observation-Based Coaching Program: Perspectives of Teachers and Coaches. *Elementary School Journal*, 121(2), 283-310. <https://doi.org/10.1086/711070>

### 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.

---

# HOW FEATURES OF THE IMPLEMENTATION PROCESS SHAPE THE SUCCESS OF AN OBSERVATION-BASED COACHING PROGRAM

---

*Perspectives of Teachers and Coaches*

## ABSTRACT

Studies investigating the effectiveness of professional development (PD) programs have provided no conclusive findings on what exactly makes a PD effort effective. Using an observation-based coaching PD program, we explore which features in the PD implementation process facilitated or impeded teachers' meaningful engagement in and learning from the PD program. We interviewed seven PD coaches and a selection of 11 teachers involved in a 3-year PD project. We found that how the school and the coach implemented the PD project played a critical role in shaping teachers' willingness to participate in the program. We conclude that the success of a PD program depends not only on enduring effective PD features but also and especially on the PD implementation process within the context of daily school life.

Yanjuan Hu  
SOUTHWEST  
UNIVERSITY

Klaas van Veen  
UNIVERSITY OF  
GRONINGEN

---

**M**ANY studies have provided insights into the observable processes and features that make teacher professional development (PD) programs effective (Borko, 2004; Desimone, 2009; Kennedy, 2016; van Driel et al., 2012), such as content focus, active learning, coherence, sustained duration,

---

THE ELEMENTARY SCHOOL JOURNAL

Volume 121, Number 2. Published online November 9, 2020

© 2020 by The University of Chicago. All rights reserved. 0013-5984/2020/12102-0005\$10.00

and collective participation. These studies mainly follow the cause-and-effect approach and focus on demonstrating a positive link between PD programs with certain features and changes in teacher cognition, teaching behavior, or student learning outcomes (e.g., Helms-Lorenz et al., 2016; Sailors & Price, 2015). In contrast, little research is available about why these PD features were effective “so that we understand under what conditions, why, and how teachers learn” (Opfer & Pedder, 2011, p. 378). It remained untested how PD implementation may affect the success of a PD program. Studies that do identify effective features describe them vaguely, and they lack information and consensus about how to translate these features into an implementation process in the daily context of schools (Santagata & Bray, 2016).

As Kennedy (2010, p. 597) points out, teacher effectiveness is a function of not only teachers’ personal qualities but also “schedules, materials, students, institutional incursions into the classroom, and the persistent clutter of reforms that teachers must accommodate.” We assume that PD effectiveness may rely not only on effective PD design features but also on how the various interruptions are managed during the implementation process. However, many PD effect studies ignore teachers’ hectic daily working environment, which explains why the field lacks understanding of how a PD program’s effective features actually work in practice and why many promising PD interventions result in disappointing outcomes. Therefore, more research into the PD implementation process is needed to understand the conditions in which key elements of the PD system are effective (i.e., the program, the teachers, the facilitators, and the context; Borko, 2004), such as research investigating specific implementation conditions in local school contexts, the roles and qualities of the person delivering the PD program, and teachers’ motivation to participate in the PD program (Borko et al., 2010; Kennedy, 2016; Knapp, 2003; van Driel et al., 2012). Such information can also be a source of learning for developing similar PD programs in other contexts in the future.

One such effective PD intervention involves observation-based coaching, which is the subject of the current study. Observation-based coaching is increasingly considered a high-quality PD program (e.g., Desimone & Pak, 2017; Kennedy, 2016) because the use of a coach and in-class observation may possibly enhance the coherence between the PD program and the needs of individual teachers. In this PD intervention, teachers are observed in their own classrooms with the help of a structured observation tool: a previously designed and validated observation instrument called the International Comparative Analysis of Learning and Teaching (ICALT; van de Grift et al., 2014; van der Lans et al., 2018). This tool enables the observer to establish a teacher’s “zone of proximal development” (ZPD) regarding teaching skills (Vygotsky, 1978). The observation is followed by a coaching conversation and a written report a few weeks later. A trained coach conducts both the observation and coaching. Studies show that this approach is highly effective in increasing teachers’ teaching skills (Helms-Lorenz et al., 2016, 2018; Maulana et al., 2015a, 2015b; Neuman & Cunningham, 2009; Neuman & Wright, 2010; Sailors & Price, 2015). However, these studies provide no insight into how this intervention actually works, how it is implemented in schools, and how teachers perceive it. In our experience, teachers reacted with mixed feelings to this intervention. Because some found the intervention beneficial and others were critical toward it, we decided to interview the coaches and teachers involved to explore in depth how the intervention is operationalized concretely in practice, for example,

what actually happens in the coaching process and how this influences teachers and their subsequent teaching practice. The main question that guided this study is as follows: Which features of the PD implementation process facilitated or impeded teachers' engagement in and their learning from the PD intervention?

## Theoretical Background

### Problems in Studying Effective PD Programs

As stated previously, much of the research to date has provided no conclusive findings on what makes PD programs effective in the context of daily work. Though a general consensus exists regarding the core design features of effective PD programs, researchers have found that the effect of these design features cannot be replicated across studies and across contexts (e.g., Kennedy, 2016; Opfer & Pedder, 2011). Kennedy (2016) suggests the reason could be that the list of design features fails to capture the underlying theories of action. She suggests that PD program effects depend on the extent to which they are enacted via teachers' independent judgment—in other words, the extent to which the teachers' opinions of the PD goals, content, and process were considered in developing and implementing the PD program. She describes four approaches through which a PD program can be enacted: prescriptions, strategies, insights, and body of knowledge. These four approaches lie along a continuum in which enactment increasingly depends on teachers' independent judgments. Her review shows that PD programs enacted via strategies and insights (i.e., stronger focus on teachers' independent judgments) were associated with more positive effects than those with prescriptions. However, PD programs that relied solely on teachers' independent judgment were less effective than those that relied on strategies and insights, as in the body of knowledge approach (i.e., the PD program presents a set of concepts and principles in the forms of textbooks, diagrams, and lectures but may not explicitly imply any particular action). This corresponds to the argument that the intensity of the features matters (Opfer & Pedder, 2011). She further suggests that program assignment methods (mandatory or voluntary) can influence teacher motivation and ultimately the program's effectiveness, pointing to the importance of teachers' own volition in improving their practices. However, it remains unclear how the intensity of teachers' independent judgments may be adjusted by the implementation process and how teacher motivation can be elicited and sustained in this process.

A further possibility is that PD programs with comparable structural design features still vary widely in the invisible cultural features embedded in the implementation process. For example, the school context may shape the invisible cultural features that make a PD program effective. In this vein, Hargreaves and Dawe (1990) conclude that contrived collegiality versus a collaborative culture among teachers could influence the effect of peer coaching on teachers' engagement in the PD program. Whereas contrived collegiality is likely to meet with teachers' resistance or strategic compliance with the PD program, collaborative cultures foster teacher development because "collaborative cultures comprise evolutionary relationships of openness, trust, and support among teachers where they define and develop their own purposes as a community. But contrived collegiality consists of administratively contrived interactions among teachers where they meet and work to implement the curricula and instructional

strategies developed by others” (Hargreaves and Dawe, 1990, p. 227). These definitions suggest that teachers’ independent judgment is influenced by the existing school culture but may also be relevant for how the school organizes PD programs. Hargreaves and Dawe (1990, p. 238) caution that “the presence of evaluation prejudices the necessary willingness to show weakness and vulnerability in order to gain support” and that administratively imposed partnerships and mandatory assignment methods may also jeopardize the creation of collaborative culture and prevent teachers from becoming genuinely committed to the PD program. Their conclusion is in line with Kennedy (2016), who concludes that teachers’ independent judgment and motivation to learn have a critical influence on program effectiveness. Both studies point to the possibility that PD program features do not function independently of one another.

The role and qualities of the person who delivers the PD program could be another source of the invisible cultural features relevant for the PD program effectiveness. Previous study shows that coaching is more effective when teachers experience cognitive dissonance and feel challenged to reexamine their existing knowledge and teaching practices (see Kennedy, 2016; Linder, 2011; Thompson & Zeuli, 1999). This typically requires coaches to identify teachers’ vulnerabilities and areas for growth and, consequently, teachers’ self-identification of these areas. Therefore, the coaches’ spirit of sincere collegiality and willingness to change and revise their ideas could soften teacher resistance to exposing their vulnerabilities (Hargreaves & Dawe, 1990). Effective coaching requires an affirming relationship between the teachers and their coach or peers and a context of trust, availability, and credibility (Hargreaves & Elhawary, 2019). It may be counterproductive if the coach appears authoritative and carries a sense of power over teachers or communicates in a patronizing manner (Heineke, 2013). Similarly, Kennedy (2016) shows that coaches in more effective programs collaborate with teachers on lesson planning rather than evaluating how well teachers complied with the recommended practices.

### Coaching as High-Quality PD

Coaching has often been recommended as a PD design feature and is increasingly considered high-quality PD (Desimone & Pak, 2017; Kennedy, 2016; Lofthouse, 2019) because teacher learning is affected not only by a set of design features, school context, and PD facilitator but also by multiple factors pertaining to teachers themselves, such as their beliefs, prior knowledge, skills, and teaching experience (Opfer & Pedder, 2011). Programs based on coaching can provide an ideal setting to deal with such complexity, considering that coaches can increase the flexibility of the PD program to meet individual teachers’ needs. The basic assumption of coaching is that instructional effectiveness can be improved when teachers receive feedback on their functioning and perceive a gap between the intended outcomes of teaching and the outcomes actually attained (Costa & Garmston, 2002). In line with this assumption, Veenman and Denessen (2001, p. 385) define coaching as “a form of in-class support to provide teachers with feedback on their functioning and thereby stimulate their self-reflection and the self-analysis needed to improve instructional effectiveness.” Their study suggested that teachers value coaching skills such as the ability to strengthen teacher autonomy and to provide timely, concrete feedback and a goal-oriented attitude. Building on this definition, coaching in our PD program refers

to a one-on-one discussion process in which the coach provides targeted feedback and engages the teachers to reflect on a selection of their instructional behaviors based on classroom observations. We refer to this process as “observation-based coaching.”

Desimone and Pak (2017) discuss coaching as a potential factor in high-quality PD programs by elaborating on the consistencies between coaching and five key features of effective programs. Some studies identify coaching as helpful in changing teacher perceptions and practices about teaching and student learning (Neuman & Cunningham, 2009; Neuman & Wright, 2010; Sailors et al., 2014). However, other research shows that coaching has limited or no influence on teachers’ classroom practices and student achievement (Carlisle & Berebitsky, 2011; Marsh et al., 2010). Thus, no conclusive evidence can be drawn regarding the effectiveness of coaching. One possible reason for the mixed findings could be an unclear specification of the coaching process, the role of the coaches, or the conditions that make coaches effective.

### Observation Tool and ZPD

Prior literature shows that teachers must experience a sufficient amount of cognitive dissonance to disturb their existing beliefs and knowledge (Linder, 2011; Thompson & Zeuli, 1999). Similarly, Kennedy (2016) concludes that PD programs are more effective when they foster new insights by raising provocative questions that force teachers to reexamine familiar events and come to see them differently. Such cognitive dissonance may act as a catalyst for change (Opfer & Pedder, 2011). However, as Coburn (2001) similarly argues, the intensity of dissonance and confrontation should match the individual conditions of the teacher, namely, not so intense as to be frustrating but challenging enough to push the teacher forward in her or his development. To determine the level of dissonance needed, we applied the Vygotsky concept of ZPD, which is increasingly applied to teacher PD (e.g., Heineke, 2013; Kuusisaari, 2014; Warford, 2011). This zone is defined as “the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers” (Vygotsky, 1978, p. 86).

To identify the ZPD (i.e., where such dissonance can be created), we used the ICALT observation tool (van de Grift et al., 2014; see app. B, available online). The six domains in this observation tool describe teaching behaviors following a cumulative order, from simple to complex and advanced teaching behaviors. Indicators based on the observation scores can be used to formulate feedback targeted at actual teaching problems the teacher is facing and at the level the teacher feels is feasible to manage. In other words, the observation tool allows the coach to scaffold the coaching feedback in response to the teacher’s ZPD.

However, PD program effectiveness will rely on the coach and how competently the feedback is delivered and received. The nature of the feedback on ZPD renders it a form of criticism, which often has a negative connotation of judgment; thus, how criticism is delivered can be crucial in shaping teacher reactions, as “people do not always respond to criticisms with logical and calm detachment” (Hornsey et al., 2012, p. 125). Feedback of a critical nature could meet with teacher resistance and defensiveness. Therefore, effective delivery of critical feedback may rely on several factors, such as the experience and expertise of the coaches, the pedagogies that



coaches adopt, the coach's perceived credibility, the style and tone in which the feedback is delivered, and the coach's motive and genuine concern (e.g., Hornsey et al., 2012; Imants & van Veen, 2010; Leary & Terry, 2012). These factors may be another reason coaching is effective in some PD programs and less so in others, as reported by Kennedy (2016). Another challenge to delivering critical feedback effectively involves the teacher's interpretation of feedback as evidence of individual deficiency or personal weaknesses, which requires a context in which the coaches allow for changes and revisions to their own ideas to the same extent that they expect such change from the teachers (Hargreaves & Dawe, 1990). This critical openness may not always be available, especially in the directive/technical coaching approach.

Thus, a successful PD program should motivate and intellectually engage teachers, which can be influenced by various aspects including situational and implementation elements, as well as the interaction among these features. With this study, we aim to gain a deeper understanding of these features embedded in the PD implementation process. We focus on the experiences of coaches and participating teachers as the most directly responsible people who influence the quality of implementation.

## Method

This qualitative study examined the experiences of coaches and teachers participating in a longitudinal PD project aimed at improving the teaching quality of secondary schoolteachers in the Netherlands. To better understand the implementation process of this PD program, semistructured interviews were conducted with the coaches and a selection of participating teachers during the third year of this project (November 2016–April 2017). An interview protocol was developed and piloted with one coach and two participating teachers. They provided feedback on clarity and relevance of the interview questions. In the interview, participants were asked to give step-by-step descriptions of how they experienced the whole PD process and which key features in this process did or did not facilitate their participation in the PD program. Interviews were conducted in English and lasted for an average of 1 hour and 40 minutes, ranging from 43 minutes to 128 minutes.

### The PD Project Design

This PD project (2014–2017) was designed to improve the teaching practices of secondary schoolteachers from a selection of schools in the Netherlands considered weak performing by the Inspectorate of Education. It was financed by the Dutch government, and the school received 2,000€ for each teacher who participated. The project consisted of systematic classroom observation (using ICALT) and post-observation discussions (i.e., observation-based coaching). The ICALT observation tool has 32 items and 6 domains, including safe and stimulating climate, classroom management, clear and structured instructions, activating students, differentiating instruction, and teaching learning strategies. These six indicators describe teaching behaviors following a cumulative order, from simple to complex and advanced teaching behaviors. Coaches can use these six indicators to formulate initial ideas about teachers' ZPD. They can then refine the estimated teachers' ZPD in the post-observation discussion and thus

scaffold the coaching feedback in response to the teachers' ZPD. The post-observation discussions are not restricted to the six indicators in the observation list. The project used trained coaches to ensure that the PD program was effectively adapted to individual teachers' specific needs within their school context. Altogether, 518 teachers from 15 Dutch schools were involved for a period of 2–3 years. Each teacher should receive in total four visits from the PD project (once every 3–6 months over a period of about 2 years). Seven coaches were appointed by the project. Before starting the project, the coaches were carefully trained how to use the observation tool. Interrater reliabilities of the observations were calculated during the training sessions to ensure a sufficient consensus among coaches (see van de Grift et al., 2014).

### The Participants

As an exploratory step to understand the diverse factors that can influence the PD implementation, we interviewed all 7 coaches and a selection of 11 teachers participating in our project (see Table 1). The coaches (5 females, 2 males) were teacher educators from a research-intensive university in the Netherlands; two of them were also teachers in schools. Most of them had taught for many years in schools before becoming teacher educators. They also specialized in different subject areas. For feasibility purposes, coach-teacher dyads were not based on matches of subject areas. From May 2014 to November 2016, they had cumulatively completed more than 900 observations and coaching sessions (though coaches differed in the number of observations and coaching sessions; see the second column of Table 1).

The 11 teachers were from 9 different schools, with varying years of teaching experience (3–33 years) and teaching subjects (chemistry, Dutch, French, creative arts, English, biology, physical education, and physics). We opted for a purposeful sampling strategy to approach these teachers. We asked each coach to help us identify two teachers they had observed for an interview: one teacher who found the project

Table 1. Background Information of the Coaches and Teachers

Coach	Number of Observations	Teacher	Number of Observations	Observation Score Pattern	Teaching Experience (years)	Subject
Nelleke	180	Theo <sup>a</sup>	3	+	7.5	Chemistry
		Alex <sup>b</sup>	3	+	2 (intern) + 1	Dutch
Lisa	68	Inge <sup>a</sup>	3	±	14	French
Roeland	199	Linda <sup>a</sup>	1	n/a	32	Creative arts
		Chris <sup>b</sup>	3	±	32	English (5 years Dutch)
Emma	53	Daan <sup>a</sup>	2	+	24	Chemistry and biology
Ben	237	Saskia <sup>a</sup>	4	±	14	English
		Fred <sup>a</sup>	3	±	15	Physical education
Gerda	157	Jasper <sup>a</sup>	1	n/a	5	Physics
		Rosanne <sup>b</sup>	2	n/a	33	French
		Jacobine <sup>b</sup>	3	+	18	Creative arts
Anne	33	n/a	n/a	n/a	n/a	n/a

Note.— + indicates an increase in the observation scores over time; ± indicates fluctuations in the observation scores over time; n/a means no information available.

<sup>a</sup> The teacher is critical of the professional development program.

<sup>b</sup> The teacher is positive about the professional development program.



helpful and one who did not. We did so to ensure that we included diverse views on both the strengths and weaknesses of the project. We asked the coaches to help with replacements when teachers did not reply to the research invitation or could not participate due to busy schedules. In about 5 months (November 2016–April 2017), we had invited 33 teachers. Ultimately, for feasibility purposes, we were able to interview 11 teachers (6 males, 5 females), 4 of whom were positive about this PD program and 7 of whom were critical of it. We created pseudonyms for each interviewee for de-identification. The teachers differed in the number of times observed and coached. The PD programs for Linda and Jasper were interrupted due to personal reasons, but we still included them because they represented a group of teachers who were not able to complete the program. Examination of their overall observation scores revealed wide variations in how the PD program affected the teaching skills of the teachers we interviewed.

### Data Analysis

We used the ATLAS.ti 7 qualitative analysis software for content analysis of the interview transcripts (Miles & Huberman, 1994). We started with a tentative coding framework based on related literature, suggesting that the effectiveness of PD programs may depend on four key elements (i.e., PD program, PD facilitator, teacher, and context; Borko, 2004), as well as the facilitating pedagogy (i.e., prescriptions, strategy, insight, and knowledge) and teacher motivation in PD. The motivation further relates to the program assignment methods (mandatory vs. voluntary), the PD coach's expertise, and the environment of the PD "classroom" (Kennedy, 2016).

The analysis followed an iterative process and comprised several steps. We started with two interview transcripts and assigned descriptive codes to interview fragments that referred to factors that facilitated or impeded teacher learning or teacher engagement in the PD program. Next, we categorized the descriptive codes into analytical codes in consultation with the tentative coding framework. We created open codes when descriptive codes emerging from our data could not be categorized into the initial coding framework. We then repeated the first and second steps with all interviews and integrated the analytical codes to further develop the tentative coding scheme. We reviewed all interview codes and recoded them as needed. Detailed descriptions of the codes and illustrative examples can be found in the Appendix A, which includes four main categories of features related to the school context (Table A1), the coach (Table A2), the teachers (Table A3), and the PD program design (Table A4).

Next, we examined the code frequencies and the code co-occurrences to discern the general association patterns between PD program effects and the influencing features pertaining to the school context, coach, program design, and teacher. PD program effects were operationalized as the impact of the PD program design and implementation on teachers' skills, attitudes, knowledge about teaching, and their emotions and commitment to the PD program (see the first column of Table 2). The Co-Occurrence Tool and the analytic function of Network Views in ATLAS.ti can be used to identify the proximity of two codes: embedded or overlapping, or if two or more codes are applied to the same quotation (Friese, 2015). We used the Co-Occurrence Tool to identify the codes about the four PD elements (i.e., school context, coach, PD design, and teachers) that co-occurred with those about the PD program effects, and then we used

Table 2. Perceived Positive and Negative Effects Mentioned by Coaches and Teachers

Perceived Professional Development Effects on Teachers	Nelleke (Theo/Alex) <sup>a</sup>	Lisa (Inge)	Roeland (Linda/Chris)	Emma (Daan)	Ben (Saskia/Fred)	Gerda (Rosanne/Jacobine)	Anne	Percentage of Respondents ( <i>n</i> = 17)
Positive effects:								
Increased awareness of teaching problems or own teaching style (intended effect)	o(X/o) <sup>b</sup>	X(o)	o(X/X)	X(X)	o(X/o)	o(X/X)	X	59
Solving existing teaching problems (intended effect)	X(o/X)		o(X/X)	X(X)	X(X/X)	o(X/X)		65
Attitude open to learning, more reflective	X(X/X)		X(o/X)	o(X)	o(X/X)	o(X/X)		59
Increased self-confidence, happy, human attention	X(X/X)	X(X)	o(X/X)	X(X)	o(o/X)	X(X/X)	X	82
Knowledge expansion				X(o)	o(X/o)			12
Reflective mind/imaginary coach	o(o/X)			o(X)		o(X/X)		24
Transformed teaching	X(o/o)				o(o/X)	o(o/X)		18
Negative effects:								
Decreased motivation, pressure, disappointment, collapse of self-esteem, frustration	o(X/o)			o(X)	o(X/X)	X(o/o)		41
Teacher resistance, emotional wedge	X(X/X)		o(X/X)	o(X)	o(X/o)	X(o/o)		47
Clash between teachers			X(o/X)	o(X)				6
Little value	o(X/o)	o(X)			o(X/o)			18

Note.—The sign o or X within the parentheses indicates whether a code is mentioned by the teachers in the corresponding coach-teacher dyad. This overview includes 7 coaches and 10 teachers because Jasper (teacher) did not mention any effects of this professional development program on him.

<sup>a</sup> Name of the coach is placed before the parentheses; names of the teachers with this coach are placed within the parentheses.

<sup>b</sup> The sign o or X before the parentheses indicates whether a code is mentioned by the coach.

the Network View to represent the associations graphically. In interpreting the meaning of the co-occurrences identified, we consulted the quotations in which these codes were co-occurring and asked what the co-occurrence was telling us about how the PD implementation process influenced teacher development (Contreras, 2011).

After we added positive or negative tags to codes that the interviewees implied as facilitating or impeding their learning, we finally categorized the relevant features embedded in the four key PD elements as a set of situational features and features of the implementation process. “Situational features” refers to a set of features or pre-conditions brought to the PD program, including PD design features; existing school climate and leadership; the coaches’ personality, experience, and expertise; and the teacher’s gender, age, entry teaching skills, beliefs, and personality. “Implementation features” refers to adaptable features during the implementation process, including how the school assigned the PD program to teachers, used the program results, created a climate to discuss the PD program, and scheduled time for the program, as well as how the coach created the coaching culture and adopted the coaching pedagogy.

## Results

We first present the variations in the perceived PD program effects and then give a general overview of the influencing factors in relation to the PD implementation. Finally, in the main body of this section we provide a detailed illustration of how the features of the implementation facilitated or impeded teacher learning and teacher engagement in the PD program, including four additional sections about the school context, the PD coach, the teacher, and the project design.

### Variations in PD Effects

Analyses of the interviews revealed considerable variations in the perceived PD program effects on the selected teachers in this study (see Table 2). The teachers experienced both positive and negative effects at the same time; some teachers felt the positive impact more strongly, and others felt the negative impact more strongly.

Nine out of the 11 teachers interviewed (see the last column of Table 2) found the project to have increased their awareness of how they taught and were able to improve slightly on differentiating and activating students, which were the intended effects of the PD program design. Both the coaches and teachers also mentioned a few other effects, such as strengthened teacher confidence, increase in reflectiveness, and more openness to experimenting with new ways of teaching. One teacher expressed, “For me, it has helped me to become even more open and reflective on the things I do on a daily basis” (Jacobine).

Despite these positive effects, some teachers also found the PD program to be “pretty useless” (Theo) or noted that they did not really change their lessons afterward (Inge and Saskia). Some teachers even experienced negative effects, such as decreased motivation to improve teaching (Theo), too much pressure to perform well (Linda and Chris), strong disappointment, and even a collapse of self-esteem due to low scores (Daan, Saskia, and Fred).

Influencing Features in General

To understand why teachers reacted with mixed feelings to our PD intervention, we found the perceived effectiveness of classroom observation-based coaching was mainly influenced by a set of implementation features in the school context and the coaching process, which further interacted with the situational features of the program design, the existing school climate, the expertise of the coaches, and teacher backgrounds.

The perceptions of the observation-based coaching by 11 (out of 518) teachers and 7 coaches suggest possible interactions between the four PD elements. We first examined the code co-occurrences to gain a general understanding of how the perceived positive and negative program effects were related to the four main PD elements (i.e., the coach, the program design, the school context, and the teachers; see Fig. 1). Figure 1 demonstrates that the perceived positive effects were more strongly linked with the coach’s characteristics than with the other three elements, indicating the importance of the coach in a PD program’s success. The perceived negative effects were approximately equally related to features of all four elements; this finding pointed to the complexity of the impeding forces, considering that ineffectiveness could arise from non-supportive features of all four PD program elements. The program design features, specifically, the observation tool, were equally relevant for both the perceived positive and negative effects, suggesting that the program design features alone cannot predict PD program effects.

In the following, we dive more deeply into the experiences and perceptions of these 11 teachers and 7 coaches to understand how the features of the 4 key elements

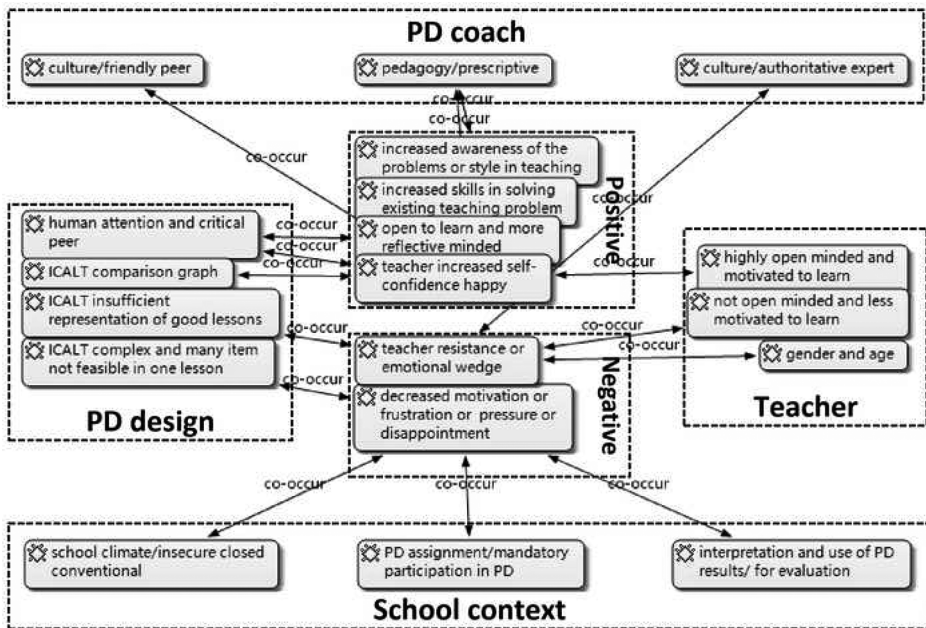


Figure 1. Co-occurrence network view of PD effects in relation to key features of the PD design, school context, PD coach, and teacher. PD = professional development; ICALT = International Comparative Analysis of Learning and Teaching. Color version available as an online enhancement.

facilitated or impeded teacher learning and teacher engagement in the PD program, as well as how some of these features worked collectively. The following results included four sections about the school context, the PD coach, the teacher, and the project design in alignment with Figure 2.

The School Context

We observed that how the school implemented the PD program influenced the extent to which teachers actively participated in it, adopted strategic compliances, or even resisted participation explicitly. The school’s motives (internal motive to improve teachers vs. external motive to demonstrate school improvement) often manifested in different PD program implementations, as reflected in assignment methods, the school’s interpretation and use of program results, the school climate, and the school’s PD program organization. These features were often interrelated.

**Assignment methods: Voluntary versus mandatory participation.** One coach and eight teachers explicitly mentioned that teachers were more active and cooperative when the school encouraged voluntary participation before, during, and after the

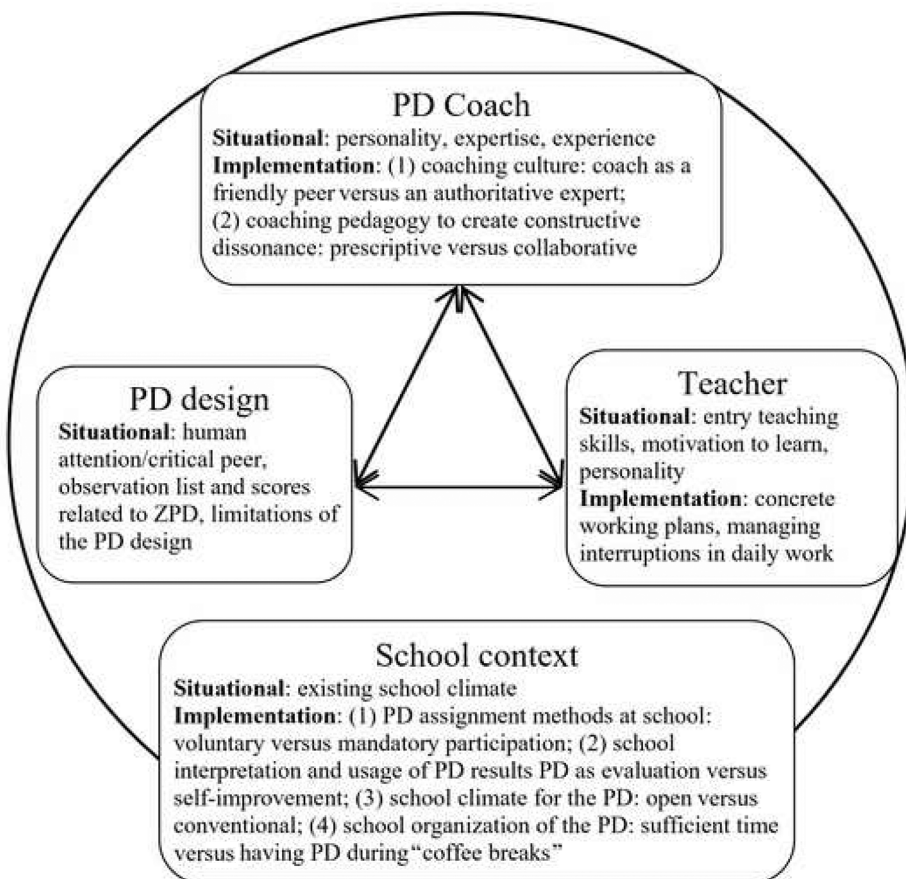


Figure 2. Relevant situational features and PD implementation features of the main PD elements. PD = professional development; ZPD = zone of proximal development.

PD program. This encouragement was evident in the school organizing introductory meetings to inform teachers beforehand about the rationale and participation process so that teachers could be persuaded and willingly participate in the PD program. The school provided the PD program to teachers as an option; as Alex said, “My team manager asked me if I wanted to participate in this project. . . . I could have said, ‘No, I’m not interested.’” Voluntary participation can be seen in the process when the scheduling was done in a joint effort with the teacher: “She mailed me and asked if she could observe one of my lessons, and then I mailed her back giving her a couple of dates and times.”

However, in some cases, mandatory participation could transform into voluntary participation if the teachers felt that they were responsible for improving their own school. This scenario was more likely in a collaborative school climate, in which the school board stressed the alliance between the school and the teachers and that the opposing pressure was from outside the school: “By that time our school wasn’t doing too well, and I felt responsible to be one of the people who are teaching quite a long time here . . . to investigate what is going well, what can go better, how can I [contribute] to the improvement of the quality of this school” (Jacobine).

In contrast, teacher resistance often emerged when mandatory participation was in place, as mentioned by three coaches and six teachers. Mandatory participation can also be reflected in every step of the process, starting from how teachers were recruited in the project. Chris had “the impression . . . that it was a bit of a top-down installment of this project. . . . It was told to us that we would participate in this project. . . . You can’t run!” Teachers felt little autonomy in the process; one teacher (Daan) suggested the project could have been strengthened if teachers were asked in advance what they would like to improve. Another teacher (Theo) disagreed with how the observation scheduling was made with him: “Management sends an email . . . ‘Nelleke has to come, there are multiple options, but we prefer Tuesday, so, Tuesday, this and that moment, she will be there. Is it okay?’ . . . It didn’t feel right for me . . . if someone else determines when and how the observation is going to take place.”

**School interpretation and usage of PD results: PD as evaluation versus self-improvement.** Closely related to the way school recruited teachers was the school’s interpretation and use of PD program results. Coach Roeland and teachers Theo, Alex, and Jacobine stressed the importance of avoiding PD results as evaluation. Teachers were more welcoming of the PD program when the school management team did not associate program results with teacher performance (e.g., not asking for the observation scores either directly or implicitly, leaving the scores between the teacher and the coach, and using scores as directions for improving teaching).

Teachers appreciated the PD program less when the school management requested explicitly or implicitly the observation results and associated the results with teaching quality and teacher performance. Two coaches (Roeland and Gerda) and four teachers (Theo, Linda, Chris, and Daan) mentioned that this was the case.

Coach Roeland echoed this sentiment, noting the director of a school asked for the observation scores and indicated that if teachers do not show the observation scores, “it could have consequences for their functioning in the school.”

The PD program as evaluation could also take the form of school boards using the observation score to show the Ministry of Education that they were no longer weak-performing schools. This elicited considerable teacher resistance, though mostly in



the form of strategic compliance. As one coach (Gerda) noted, the teachers focused more on getting higher scores by just preparing the lesson before each observation, but they soon returned to their normal practice when the coach was gone. Another teacher (Daan) was shocked by the low scores of his first observation and even explicitly said the following: “I thought it was a good lesson, because I did what I wanted to, and I got my goals. For the second observation, I said to myself, ‘I’m going to show them that I can play with the system,’ and I knew she was coming. ‘You want to see this? You will see this.’ Then I had 33 of 34. So, it’s a way of playing.”

When teachers perceived the PD results as evaluation, they became mainly focused on the quality of the evaluation, often feeling that they were being inaccurately evaluated. They thus heavily criticized the limitations of the PD program design and reported experiencing pressure and frustrations. One teacher (Daan) recalled “a clash between the high and low scorers” in his school. Although acknowledging that the project helped him improve some teaching behaviors, he bitterly concluded that his self-esteem suffered as a result of the low but misrepresentative scores. Due to such perceived evaluation, teachers tended to see the PD project as a covert threat and became critical of it. This teacher and his school showcase how PD results as evaluation can lead to ineffectiveness and even negative effects on teacher emotions and motivations.

**School climate.** Another often-mentioned, helpful feature was a school’s open, supportive, and safe climate, which depended on how the existing school climate was translated into the PD implementation climate at school. Six of the coaches (except Gerda) and seven of the teachers (except Alex, Inge, Linda, and Jasper) stressed this aspect. In schools with such a climate, mistakes were better tolerated and normalized. Teachers felt more secure being observed, and they tended to visit and discuss the lessons with colleagues for further reflection and improving their teaching, as shown in the following: “Everyone was fresh, everyone wanted to improve themselves, and it’s still alive . . . and I was curious. I’m not afraid of someone in my class, I’m not afraid of someone who is watching [and thinking] ‘What are you doing wrong?’” (Rosanne).

In contrast, in schools where the overall atmosphere encouraged behaving within the realm of the ordinary (i.e., trying out new ideas and expressing alternative opinions were considered unconventional), it was emotionally threatening to have a visitor in the classroom; thus, the PD program prompted a considerable amount of teacher resistance. Six of the coaches (except Anne) and three of the teachers (Theo, Linda, and Saskia) brought up this point. Especially in schools with mandatory participation in our project, teachers “denied” the coach’s observation about the lesson or tended to refuse to cooperate (Coach Nelleke). As Coach Roeland mentioned: “I had one teacher. . . . The lesson wasn’t good. And I said, ‘We will talk about it, are you joining me?’ He said, ‘First, I want to smoke a cigarette outside.’ I said, ‘Okay, I’m waiting in the room.’ . . . I never saw him again. He disappeared!”

Another aspect involved a lack of transparency in communication. For example, if a school’s external motivation to participate in the PD project was communicated to teachers as an internal dedication to improve its teachers, but in reality the school encouraged teacher participation because of the money the school received for each participating teacher, teachers could eventually learn about the inconsistency. Theo was clearly demotivated when he learned that his school obtained money for his participation but never mentioned it to him. He was even more frustrated when the



school management declined to use the money for his further PD, causing him to conclude that his school simply did not care about teacher improvement.

**School organization of time and space for the PD program.** Both teachers and coaches mentioned that sufficient time and smooth scheduling of the observation and coaching discussions were helpful. Some schools took care to schedule a 1-hour discussion time for each observed lesson on the same day. Failure to do so interrupted the PD implementation process. Coach Gerda mentioned this was the case in some schools: “Sometimes they said, ‘Okay, maybe you can do the discussion during the coffee break, or during the lunch break.’ And then the teacher said, ‘No, I don’t have time, because I have to pick up my children from school, so maybe we can do it next time?’ I did that a few times, a few weeks afterwards, but that’s quite strange to do.”

### The PD Coach

The successful implementation of this PD program was also relevant to the way coaches delivered the PD program. The analysis suggested that to be effective, a coach must be a “critical friend” to the teachers, as exemplified by Alex’s comment: “It had to be confronting, in a nice way of course.”

**Coaching culture: Becoming a “friend.”** All seven coaches stressed that an essential first step is to build a harmonious relationship with teachers. As revealed in the analytic coding of the interview, we found that the coaches used several specific strategies to do this, including the following:

1. Portraying themselves as peers of the teachers, being ready to revise their ideas and suggestions, and avoiding the impression of a superior who carries a sense of evaluation;
2. Creating a warm and sincere atmosphere by attentive listening; giving compliments; showing respect, trust, and enthusiasm; caring about teacher emotions and concerns; and investing in human connection; and
3. Ensuring a safe and confidential climate by stressing that the results and discussions are between the coach and the teacher and will not be shared with the school.

This rapport made it easier for the teachers to express their concerns and to accept the coaches’ feedback and suggestions. The teachers needed to feel the coaches were “one of us.” Thus, coaches found it helpful to stress their teaching experience at school, especially for some teachers who initially had the impression that the coach was a researcher from the university who had no idea what teaching was like.

Two of the teachers emphasized that the coach is someone with “an open, warm personality that gives comfort to the teacher. . . . You don’t burn him if he does things that are not on the list” (Fred) and “[who] is willing to discuss, openly, without judging” (Linda). This was especially important as mentioned by teachers from schools with relatively closed cultures and from schools that requested mandatory participation.

**Coaching pedagogy: Creating constructive dissonance.** In addition to an open and collaborative coaching culture, the teachers also found it helpful to experience constructive cognitive dissonance, which depended on the coaching pedagogy. The pedagogy used by our coaches can be roughly categorized into prescriptive pedagogy and collaborative pedagogy. Both were found effective in creating cognitive dissonance in

teachers if they were targeted correctly at the observed teacher's ZPD and challenged them sufficiently on that zone. This is because teachers expect to be intellectually challenged and want the coaches to be "somebody who is capable of looking at the problem from different perspectives, capable of interpreting what I am saying and then asking the [right] question" (Theo).

Prescriptive pedagogy was helpful in raising teacher awareness of and in solving an actual teaching problem in their current teaching only if the coach had recognized the teacher's ZPD. The ability to sufficiently recognize the zone associated closely with the expertise of the coach. The use of prescriptive pedagogy also required deliberate effort on the part of the coach to establish an open and collaborative coaching culture. Another example illustrated the importance of prescriptive feedback to be targeted at a problem a teacher was aware of but unable to solve: "That was something Ben pointed out to me, which I thought, 'Yeah, I already knew that,' and Ben said: 'You actually have to do it. Just give them [the students] a task beforehand: "this is what you need to do, and while you're doing that, I'm going to check your homework.'" That was a very helpful thing. Although I was aware what was going on, I still didn't change it until he said, 'Well, there are easier ways of doing that'" (Saskia).

Collaborative pedagogy was found to be more useful for the skilled and motivated teachers to expand their current knowledge and skills, to develop a reflective mind, and to transform their teaching. Rapport is often established by teachers in joint efforts. In this approach, coaches often withheld their opinions of certain observed teaching activities, provided script tapings of specific classroom events, probed for evidence of teachers' claims, engaged teachers in joint inquiries about how to improve a lesson, and elicited teachers' thinking by raising proactive questions that forced teachers to reexamine familiar events and come to see them differently. However, some form of the collaborative approach might create too little cognitive dissonance, especially when the coach provided only positive comments, which might be perceived as unreliable and even be mistaken as patronizing. In one case, an experienced and skilled teacher (Inge) concluded, "If a person says to you, 'Well, your lesson was very good,' then it's not really helpful. It's [only] cheering you up! . . . There was no feedback, actually, to improve myself. It was just teeny tiny little things."

### The Teacher

Though the school context and the coach could have a critical influence on the successful implementation of this PD program, teachers' own motivations, personalities, and entry teaching skills also played a role. Six of the coaches and all 11 teachers mentioned this factor. Some teachers volunteered and were eager to improve their teaching; some were already good at teaching, slightly open for learning new things, but not really having great expectations; and a few were just there to "earn their bread" (Roeland). These teachers also reacted differently to the PD intervention, which led to different effects.

To learn from the PD intervention, Jacobine, a teacher who became more open to learning after participating in the project, said, "Teachers should be a little bit vulnerable, you should be open, you should be curious, reflective, and not afraid of being criticized." Teachers who were already good at teaching when entering the project were most likely to evince these traits. Teacher resilience can be another helpful feature;

one teacher who was strongly disappointed at the observation scores from the first observation said, “I was too proud to give in, so I have to learn. I have to let them see I’m better than 19 points” (Daan).

Another relevant factor was how teachers themselves managed their participation in the PD program against some opposing forces in their school context. For example, two of the teachers were motivated to learn and had good connection with the coach; one (Theo) was distracted by the perceived impeding factors from the school (e.g., mandatory participation, PD results as evaluation, and nontransparent communication), and the other (Linda) managed to separate the influence from the school and focused on the positive outcomes from the project.

Related to the motivation and personality of teachers, the teachers who made concrete working plans found the project helpful. Some teachers were eager to improve their teaching, intentionally making and following working plans after the observation and discussion. One teacher (Alex) in the interview said that he had imagined that there was a coach in his classroom during the lessons when the PD coach was not present. He said he tried this imaginary coach idea for 4–5 weeks until his attention shifted away. This extra intentional step actually created more observation and discussion sessions, which were missing for many other teachers in the project. This example also points to the need to motivate the teachers so that they become willing to devote additional efforts to compensate for the limitations of the PD design.

### The Project Design

The teachers and coaches highly valued the PD program because the on-site, observation-based coaching design provided a formal moment and a critical peer for teachers to reflect on their lessons, which was largely missing in the everyday teaching practices of experienced teachers. The PD design also had some limitations, as perceived by our participants, which have negatively influenced teacher participation in the PD program. Just the negative influences deriving from the PD program limitations were only found when the school used the PD program results for evaluation purposes or when the coach failed to approach them with an open and nonjudgmental mindset. In other words, the features in implementation by the school and the coach interacted with the design features of this PD program. Different interaction patterns led to different effects on teachers as illustrated below.

**Human attention and critical peers.** Although the main purpose of this project was to improve teaching behaviors through systematic observation and post-observation discussion, the project was actually more valuable for providing much needed attention and a moment for teachers to reflect on their everyday teaching practice. Five coaches (Nelleke, Emma, Ben, Gerda, and Anne) and eight teachers (Theo, Linda, Chris, Daan, Saskia, Jasper, Rosanne, and Jacobine) explicitly mentioned this as the most valuable outcome of their participation in the PD program. One of the coaches, Anne, noted, “What I think the power of this project is that people get attention.” In the words of a teacher: “When I’m a teacher, I close the door, this is my house, I am the boss, I do it my way. . . . You never get someone telling you, ‘Hey, why don’t you do this?’ . . . In that way, it’s a very lonely way of living. You only have your students [and] . . . they’re not very critical of the lesson . . . that’s not how we work at school” (Daan).

Coach Nelleke echoed this reaction, saying, “The positive reactions are because they feel that I took the time to listen to them; they are given compliments. They are not often seen by the headmaster or the board of the school, it is more a human interest.” However, teachers also cautioned that this human attention could be interpreted as a threat to their current teaching practices if the school used PD results as an evaluation (Linda, Chris, and Daan) and if the coach approached the teacher with a sense of superiority (Anne and Alex).

**Observation list and scores.** The observation list and the observation scores had both positive and negative influence over teachers. Some improved teaching practice because the list provided consistent observations over time and was clear and concrete for reflections afterward (coaches Lisa, Ben, and Anne; teachers Theo, Alex, Linda, and Jacobine).

However, others were demotivated due to limitations of the list (coach Nelleke; teachers Theo, Daan, Saskia, and Fred). The limitations included the observation list not capturing some important aspects of good teaching and not being applicable to practice-oriented lessons. For instance, two passionate, confident, and experienced teachers (Daan and Fred) found the list to have raised their awareness of where they needed to improve, but at the same time the misrepresentative low scores were demotivating.

To understand why some teachers found the list and scores constructive and other teachers were demotivated, we found that the teachers from the former group willingly participated in the PD program (i.e., Alex, Rosanne, and Jacobine), whereas the latter mostly felt pressure to participate (Theo, Daan, and Fred).

**Limitations of the PD design.** Some of the coaches and teachers mentioned reduced teachers’ commitment to the PD program due to a few limitations in the program design. First, the observation and coaching discussions took place only twice a year, which was too infrequent and random (Nelleke, Roeland, Gerda, Alex, Inge, Saskia, and Jasper): “It can be useful in that moment, but you forget about it if there’s no continuation” (Jasper). Second, the feedback and observation results were not accurate enough, as the coaches only visited one lesson each time (Theo, Daan, Saskia). Third, as Alex and Daan noted, the program design involved little teacher independent judgment, as the teachers were not asked beforehand what they would like to improve or focus on but were all evaluated on a fixed predesigned list. However, like the mixed effects of the observation list, these limitations only hindered commitment to the PD program when teachers perceived that their participation was mandatory and that the results would be used as an evaluation of their teaching performance, as in the case of Theo and Daan. In contrast, Alex’s school encouraged voluntary participation; though he mentioned these limitations, he still found the PD program to have helped him to become more reflective about his teaching.

## Conclusion and Discussion

We investigated factors that influence the implementation process and, consequently, the success (i.e., teacher learning and teacher engagement) of an observation-based coaching PD program. These factors, as summarized in Figure 2, pertained to a set of features including the school context, the PD coach, the teacher, and the PD design. Successful implementation of this intervention required dynamic, interactive relationships among these four main factors.

Specifically, we found meaningful teacher engagement in the PD program typically occurred in three interrelated conditions: (1) voluntary participation in the PD program, (2) a safe and collaborative PD culture, which allowed the dissonance to be constructive rather than destructive, and (3) the creation of sufficient dissonance between what teachers already know and the new information from the PD program (cognitive/conceptual friction). These conditions were embedded in the PD design and how the school and the coach implemented the PD program. The conditions relating to voluntary participation and the PD culture were further embedded in the school context (e.g., how school management introduced the project to teachers and used the observation scores and how typical it was to ask for help) and the coaching culture (e.g., the extent to which a coach was critically open to disagreement of the teachers, connected with teachers as peers rather than authorities, and created a sense of closeness and alliance rather than a flat and aloof business relationship). The dissonance was reflected in the observation tool and the pedagogy of the coaches, mainly related to the coach's ability to adequately recognize the teachers' ZPD and provide practical suggestions accordingly. Failure to meet these three conditions could impede teachers' willingness to participate in the PD program. Other related features included the teacher's personality and motivation to learn and some features of the PD program design, which can modify the creation of an effective PD climate and constructive dissonance, eventually influencing learning.

Therefore, we contribute to current PD research by describing a set of interrelated conditions that empowers or impedes the successful implementation of the observation-based coaching program. We also contribute by showing that teachers' meaningful engagement in a PD program results from the interplay of program design features and cultural aspects (e.g., school climate and coaching culture) in the implementation process.

### Teacher Cognitive Dissonance and PD Culture

Examining the implementation process revealed that PD program success relied on finding an optimal intensity of dissonance and establishing a supportive PD culture. The coach can create this dissonance during the coaching conversations and with the observation scores in the written report. Previous studies note that teachers must experience a sufficient degree of cognitive friction to start reflecting on their teaching (e.g., Kennedy, 2016; Linder, 2011), though others caution that too much dissonance could result in teachers' dismissal or rejection of new ideas (e.g., Coburn, 2001; Opfer & Pedder, 2011). However, understanding is limited regarding how to ensure that the dissonance falls within individual teachers' ZPD. The current study shows that one way is to focus on a problem a teacher is partly aware of but not yet able to solve (e.g., Ben and Saskia). The coach can then adjust the intensity of the dissonance by, for example, providing practical suggestions or script taping (i.e., quoting the exact sentences teachers and students said during a lesson) that target the problem, according to the extent of teacher awareness of the problem and teacher resourcefulness.

We found that an open and collaborative PD culture was a prerequisite for PD program success. The PD culture was a mixed result of the school's culture for talking about program results and how the coach established a rapport with teachers. Other researchers report that a collaborative culture is necessary to reduce teacher resistance and defensiveness when receiving critical feedback (Hargreaves & Dawe,

1990; Hargreaves & Elhawary, 2019; Hornsey et al., 2012). For example, Hargreaves and Dawe (1990) describe how the collaborative culture (trust and sharing) may turn into contrived collegiality in the presence of evaluation or administratively imposed partnerships. Relatedly, Kennedy (2016) stresses the need to acknowledge the role of motivation in PD. Our study provides a detailed description of the major influence on the PD culture and teacher motivation of how the school implements a PD program. The school implementation was often a manifestation of a school's motivation (i.e., whether it was internal, to improve teachers, or external, to show school improvement). This motivation was further reflected in the program assignment methods (voluntary vs. mandatory), in how the school used program results (as evaluation or for teachers' self-improvement), the PD climate (safe/open vs. insecure/closed), and organization of time for the PD intervention (1 hour for coaching vs. coaching during coffee breaks). The teachers valued the PD program more when their schools demonstrated internal motivation, represented by giving teachers an opportunity to opt out of the PD program, not requesting the PD program results, valuing teachers' talking about teaching problems and asking for help, and organizing sufficient time for the PD intervention.

How the coaches implemented the PD program also powerfully influenced the PD culture and the creation of cognitive dissonance, as manifested in the coaching culture they created and the coaching pedagogy they used. Coach identity, personality, gender, age, experience and expertise, and tone of communication can all play a part in shaping the coaching culture. Similar to the school culture, a collaborative coaching culture seems to be a precondition for any coaching pedagogy to be effective. Collaborative dialogues were helpful for teachers to expand their knowledge base about teaching, and this approach often worked better with teachers who were motivated and competent. Prescriptive suggestions helped raise teachers' awareness and solve problems in their current teaching. It can be more challenging for coaches to provide effective prescriptive suggestions, as they require a successful establishment of an open coaching culture and the skill to challenge teachers in their ZPD. Our results are comparable to previous findings on the effects of directive coaching, which can be effective when provided in a context of critical openness (Hargreaves & Dawe, 1990; Sailors & Price, 2015). They also correspond with previous findings on effective coaching discourse, in which Heineke (2013) shows that healthy coaching relationships are incompatible with the presence of authority and power between coaches and teachers. Effective coaching relationships should be safe, confidential, nonevaluative, and non-patronizing. Our study also shows that coaches were effective when they appeared positive and encouraging, demonstrated a genuine dedication to teacher improvement, deliberately avoided the authority/expert impression, and stressed that they were the teachers' peers. These features can be explained by the "intergroup sensitivity effect" (Hornsey, Oppes, & Svensson, 2002, p. 293), in which people better receive critical feedback when it comes from an in-group member who is genuinely motivated to create a positive change instead of trying to reinforce his or her own superiority.

### Relationship between PD Design Features and Implementation Features

In line with Opfer and Pedder (2011), who note that PD program features can work collectively together in different ways, we found the implementation features were closely related to a set of preconditioned and situational features, such as the existing



school climate (related to the PD climate), the expertise and personality of the coach (related to the coaching culture and pedagogy), teacher skill and personality, and PD program design features. These results indicate that future PD studies should consider how coaches were selected and trained and how teachers were oriented before entering the PD program.

We also found that PD program design features work collectively with PD cultural features, especially the limitations of PD program design with PD culture at school and the coaching culture. Teachers were able to focus on the strengths of the PD program and managed to use the program to improve their teaching when the school gave teachers a free choice to participate and remained unaware of the program results. In particular, if schools could avoid using program results as evaluation, teachers likely could tolerate the limitations of the program design and focus on its positive value to reflect and improve their teaching. Teachers appreciated the human attention and concreteness of the observation list when the coach succeeded in relating to teachers as a friend who had a critical eye toward their ZPD. In contrast, teachers heavily criticized the limitations of the PD program design when the school used mandatory participation and associated program results with teacher performance. Therefore, the interactions between different features can influence a PD program's resilience and eventually influence the extent of the program's effectiveness.

In addition, program resilience was reflected in the coach's flexibility and teachers' self-reflection and execution ability. It was not always possible for a teacher to simultaneously experience all the effective features of the school context, the coach, and the teachers. For example, teachers in our study often mentioned a lack of effective school implementation features, such as when the school used the observation results as an explicit or implicit evaluation or was motivated by the incentives provided by the PD project but failed to show dedication to actually improving the school and its teachers. However, in such a situation, teachers may still learn from the project if the coach portrayed the PD program as an opportunity by showing genuine concern for positive changes and did not interpret observation results as an evaluation. In the face of teacher criticism of the PD program design limitations, the coach needed to explain the limitations of the design and be more careful in using the instrument adaptively as an assistive tool for discussion. Furthermore, if the features pertaining to the design were ineffective (e.g., lack of duration and follow-up), teachers could compensate by inventing additional observations and coaching themselves. The teacher who invented the imaginary coach was an example of this response.

We note some limitations of the PD program design features. The observation-based coaching design seemed intrusive for most teachers, and the scores were confrontational. The use of a predesigned list made the PD program strongly prescriptive in nature. We recommend that teachers be consulted about the observation list beforehand when designing future PD programs. Furthermore, the 6-month intervals of the intervention were too infrequent. The PD program could also have been strengthened if working plans were followed. Nevertheless, these design features can still be interpreted as both limitations and strengths, depending on how the school and the coach implemented the PD program, which points to the need to strengthen our capacity regarding managing program design limitations in the implementation process. However, these perceived design limitations may be unique for the large-scale PD program in our study, which could have limited implication for smaller-scale



PD programs and those that incorporate research- and evidence-based practices for effective PD. We also found that the PD program was valuable to teachers simply because it allowed for formal time to reflect on their lessons and a critical peer to provide feedback, which were largely missing in the everyday teaching practices of experienced teachers.

However, we should note, for instance, the data collected by interviews may have limitations in characterizing the school climate and the coaching culture. We should also note that our conclusions are based on the perceptions in a retrospective view rather than on the coaching experiences at the time. Such focus on participants' perceptions may have limitations for the understanding of PD program effectiveness and implementation. Although we took care to ensure neutrality and objectivity, it is likely that how the teachers reported their experiences may be influenced by their satisfaction with the PD program implementation. We recommend further research to include both observational data and interviews to gain a fuller understanding of the implementation process. Also, we have chosen to include a small sample in this study as an exploratory step to understand the factors that influence program implementation. Thus, our findings should be further validated when larger sample sizes are available.

In closing, the observation-based coaching PD program we investigated influenced teacher commitment to the PD program and teacher learning in various aspects and with varying degrees of impact. How the school and the coach implemented the PD program played critical roles in shaping its effectiveness. Our PD program was most effective when the school and the coach jointly created a safe, nonevaluative PD culture and when the coaches successfully created cognitive dissonance near the teachers' ZPD. We provide a rich description of the relevant features that influence the implementation process. Although ideally PD programs are most effective when implemented with all effective features, it may not always be possible in the actual implementation. We therefore introduce the question of PD resilience for developing valuable PD opportunities for teachers. As Kennedy (2010) stresses, teacher effectiveness is the combined function of personal qualities and situational forces; similarly, we conclude that PD program effectiveness is a function of not only enduring effective PD program features but also various interruptions along the implementation process within daily school life. The most powerful force in this PD intervention, and perhaps all interventions, is the aspect of human and professional attention.

## Appendix A

Table A1. Coding Scheme and Examples of Influencing Factors Related to School Context Identified by Professional Development (PD) Coaches and Teachers

Category	Definition	Example
PD assignment methods:		
Mandatory participation (–)	This code is used when the participants mentioned whether teachers were given a choice to participate or to say what to be observed or when to be observed.	I think there was too much pressure that you had to do it. And of course, if you would say, well, I don't like that, you get a wrong impression with your management, of course. "Why don't you want to do it? Why don't you? Do you have something to hide?" (Chris)

Table A1 (Continued)

Category	Definition	Example
Voluntary participation (+)	This code is used when teachers' independent judgment was consulted when recruiting them to the PD program or during the process (e.g., the school introduced the PD rationale and participation process to teachers; teachers were consulted whether they would like to participate; teachers could refuse to participate without visible and invisible sanctions).	I see it more as an opportunity, yeah. I could have said, "No, I'm not interested."(Alex)
School interpretation and use of PD program results	This code refers to how the school handled the PD program results at school (e.g., whether the results were known to the school manager; the school either explicitly or implicitly associated teacher International Comparative Analysis of Learning and Teaching (ICALT) results with teaching performance, to show improvement to the Ministry of Education, to compare with other schools).	But still, because in my experience, because we're at a school where, at that time, there was an authoritative leader, you still felt like you had to do well. And if you talk about freedom, if you don't feel like you can fail, you feel differently about visitors. (Linda)
School climate: Safe, open, collaborative climate	This code refers to an environment in the school that fosters teachers feeling they can work together, discuss their teaching problems, and help each other. The school leader believes that teachers keep learning and encourages teachers to visit and discuss lessons with each other. There is a shared sense of trust and transparent communication. Teachers have the idea that they can disagree with the boss.	The school environment is very important. If the school board says it is alright to make mistakes as long as you learn from it. Compared with a school where you are not allowed to make mistakes otherwise you have to go. I think for teachers to learn there has to be safe climate, in which you feel it is alright to make mistakes. Just be open for mistakes, ask for help and learn from it! That's the climate we need. (Lisa)
Insecure, closed, conventional climate	This code refers to lack of openness to critiques among colleagues and with the school manager where teachers feel stress, threat, tension, and even sanction when they express disagreement with the boss who appears to be an authoritarian leader. Colleagues show a tendency to act within the realm of the ordinary. Trying out new ideas and expressing alternative opinions are considered unconventional.	We had someone in the management who I found very . . . high authority? He was like a father who says all the things you did wrong, not the things you did right . . . so the things you did wrong were commented on, but the things you did well were not observed . . . because he didn't let space to have a difference. (Linda)
PD organization	This code refers to school organization of time and space for observation and coaching.	At the small school, XX, it was organized in a very good way. Always when I came there, every lesson that was planned took place, and an interview afterward was always possible. (Gerda)

Table A2. Coding Scheme and Examples of Influencing Factors Related to the Coach Identified by the Professional Development (PD) Coaches and Teachers

Category	Definition	Example
Coaching pedagogy:		
Prescriptive pedagogy	<p>This code refers to coaching pedagogy that is relatively low on teacher independent judgment and is more regulated by the coach.</p> <p>Specifically, the coach raised teacher awareness of a teaching problem by explicitly telling the teacher a problem as observed, explicitly describing or demonstrating their ways to address a particular teaching problem, advised teachers what is best to do, and explained why teachers scored 2 or 4.</p>	<p>She [Nelleke] says, "I saw something going on with those students in the corner, and you didn't correct them. You have to do that if you want everybody to actively participate in your classroom. You could have asked . . . what did they learn after your class?" (Alex)</p>
Collaborative pedagogy	<p>This code refers to coaching pedagogy that is relatively high on teacher independent judgment and is more regulated by the teacher.</p> <p>Specifically, the coach engaged teachers in joint inquiry, brainstormed about how to improve a lesson, probed for evidence of teachers' claims, raised proactive questions that force teachers to reexamine familiar events and come to see them differently, and challenged or confronted teachers about a problem in their teaching.</p>	<p>[The coach] said, "Did you see that some of your students do nothing in their lesson?"</p> <p>And I [the teacher] thought, "What about doing nothing?"</p> <p>"Are they doing nothing because they're done with their homework, or are they doing nothing because they don't get it?" [The coach asked.]</p> <p>"Well," I said, "one or two are in the first group, they [did] well, but also one or two are not getting it."</p> <p>"Why don't you use them?"</p> <p>"Hmm, that's a good question. I don't know why. Let's try it." (Daan)</p>
Coaching culture:		
Friendly peer	<p>This code refers to invisible features of the coaching environment that the coaches interact with teachers as if they were one of the teachers, walk in the shoes of the teachers, are empathetic, and told teachers that they were not there to judge the quality of teaching but to offer help. Teachers viewed the coaches as "one of us."</p> <p>This code also refers to quotations when a coach was mentioned to be enthusiastic and tentative listener, showed respect and trust, gave compliments, cared about teacher emotions, and invested in human connection and building a rapport.</p>	<p>Well, it makes you feel that you are an equal partner. You're observed as someone who is serious; you can have a conversation and you're not being looked down on. (Jacobine)</p> <p>I don't have the confronting coaching style. I really like to do it, especially if you just come in the classroom and go out, you can hurt people a lot if you give the feedback in a hard way.</p> <p>"This is not good, this is not good, I haven't seen this, I haven't seen that". . . Like managers do that . . . I have seen how damaged they can be if a manager comes into the classroom for their observation, and the way they give feedback, it stays in their mind for their whole life. They are really really damaged, or think "I'm really not fit for the job." I think it's bad. (Anne)</p>

Table A2 (Continued)

Category	Definition	Example
Authoritative expert	<p>This code refers to invisible features of the coaching environment (e.g., the coach took the role of “expert” or knowledgeable specialist, was more assertive, or showed a tendency to control and evaluate in checking International Comparative Analysis of Learning and Teaching [ICALT] scores). Teachers felt the coaches did not allow revision of coaches’ ideas and treated them in an authoritative and patronizing manner.</p> <p>This coded also refers to quotations when a coach was mentioned to be relatively detached and remote, considered the coaching as a job responsibility, did not devote extra time, was not personally passionate, kept teachers at a distance, cared less about teacher emotions, or communicated explicitly the weakness of teachers.</p>	<p>I’m always asking, “Am I right [that] I didn’t see this? Am I right [that] I didn’t see that?” . . . “Is this never happening in your lessons? Or is it just this lesson? Why did you make this choice, not to show this? You knew the ICALT. So you knew what I came to score. And you’ve chosen . . .” (Ben)</p>
Situational features	<p>This code refers to some relevant features of the coach themselves such as seniority, expertise, experience, still teaching at school, open-minded, not judging.</p>	<p>I was a teacher for 20 years, so I know how children react to things. I think if you need to ask the right questions, bring up exactly the right things, the things that are really the trouble in the lesson. (Emma)</p> <p>Very open, she gave me the impression that she listened to me, that she was reacting to things I said, and that she was not just giving her opinion. (Alex)</p>

Table A3. Coding Scheme and Examples of Influencing Factors Related to the Teachers Identified by Professional Development (PD) Coaches and Teachers

Category	Definition	Example
Situational (gender, age, personality entry teaching skills, motivation to learn)	<p>This code is applied to interview fragments about teachers’ individual conditions such as entry teaching skills, whether they are critical, open-minded, curious, and reflective about their teaching.</p>	<p>He said, “I’m a teacher here to earn my bread.” He said, “Well, I’m here, it’s my job and I have to go still 20 years.” And that was his attitude. (Roeland)</p> <p>She was so good in reflecting. She was very eager to learn. But also, she was very curious about what exactly that according to me is not working. (Emma)</p>
Implementation (self-management of their participation in the PD)	<p>This code refers to whether teachers made concrete working plans and how teachers themselves managed their participation in the PD program against interruptions in daily school context.</p>	<p>I think if Nelleke [the coach] was in my classroom this lesson, what would she have said? . . . What is the thing that stands out from this lesson, what would Nelleke say about that? . . . I’m thinking after most of my lessons. (Alex)</p>

Table A4. Coding Scheme and Examples of Influencing Factors Related to the Professional Development (PD) Program Design Identified by PD Coaches and Teachers

Category	Definition	Example
Human attention and critical peer	This code is used when participants mentioned the PD program provided much-needed human attention, an extra eye, and a formal moment for teachers to reflect on teaching.	So for him, talking about it made him think about his teaching . . . finding for himself ways of “what can I do better? And I should search for that,” but he formulated at the end of the conversations things like, “Oh I could do this even better, or I could spend more time on doing this or that.” (Emma) Sometimes you’re getting lazy. And then, with the project, you’re getting a kick. (Rosanne)
Observation list and scores:		
Observation list (+)	This code is used when participants mentioned why using the observation list was helpful (e.g., when used as a starting point for discussion, the list is concrete, consistent over time, objective, and shows directions for improvement; the International Comparative Analysis of Learning and Teaching [ICALT] scores confronted teachers to realize certain issues in their teaching; high scorers get confirmed with their teaching quality).	I think you need the form. . . . Because if she only talks to me, then I, tomorrow I remember 95% and the day after tomorrow it’s 80%, and by the end of the week I only remember 50% of what she said. And if you have this report, or a report, then you can see, “oh yeah.” And now, because this was 2 years ago, I can still look it up: What did I do then? (Rosanne)
Observation list (–)	This code is used when participants mentioned why using the observation list was not helpful (e.g., when used for evaluation, ICALT list is an insufficient representation of a good lesson, not capturing the invisible features; low scorers get demotivated. The comparison graph with other teachers was demotivating).	There are so many other things that do happen in a lesson, which is not really something that you can fit on the sheet, so in that sense it doesn’t really give a good view of what lesson is a good one or not. And then it remains somewhat in the lower levels, instead of having a lesson of a higher quality. So I think that even though all those things at the end weren’t really in a lesson, the lesson could still be of fairly high quality. (Saskia)
Limitations of the PD program design	This code refers to aspects teachers and coaches mentioned as limitations of the PD program design, such as observation frequency, follow-ups, and so on.	It’s now like an incident. You’re coming once, and the next year another time. And that’s it. (Roeland)

## Note

The authors would like to thank all participants for sharing their experiences and perspectives. The authors declare that they have no conflict of interest. Yanjuan Hu is associate professor in higher education at Southwest University in China. She obtained her PhD at ICLON, Leiden University Graduate School of Teaching, and then worked as postdoc researcher in the Department of Teacher Education at the University of Groningen. Her current research focuses on how the organizational and cultural conditions influence teaching and learning practices in higher education. Klaas van Veen is professor in educational studies at the University of Groningen in the Netherlands. His

main research interest refers to the pedagogy of teacher learning in the workplace, inspired by Lora Bartlett's statement: "you're getting the learning you organize for." Correspondence may be sent to Yanjuan Hu at [huy@swu.edu.cn](mailto:huy@swu.edu.cn).

## References

- Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher*, *33*(8), 3–15.
- Borko, H., Jacobs, J., & Koellner, K. (2010). Contemporary approaches to teacher professional development. In P. L. Peterson, E. Baker, & B. McGaw (Eds.), *Third international encyclopedia of education* (pp. 548–556). Elsevier.
- Carlisle, J. F., & Berebitsky, D. (2011). Literacy coaching as a component of professional development. *Reading and Writing*, *24*(7), 773–800.
- Coburn, C. E. (2001). Collective sensemaking about reading: How teachers mediate reading policy in their professional communities. *Educational Evaluation and Policy Analysis*, *23*(2), 145–170.
- Contreras, R. B. (2011). Examining the context in qualitative analysis: The role of the co-occurrence tool in ATLAS.ti. *Newsletter*, *2011*(2), 5–6.
- Costa, A. L., & Garmston, R. J. (2002). *Cognitive coaching: A foundation for renaissance schools*. Christopher-Gordon.
- Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher*, *38*(3), 181–199.
- Desimone, L. M., & Pak, K. (2017). Instructional coaching as high-quality professional development. *Theory into Practice*, *56*(1), 3–12.
- Friese, S. (2015). *ATLAS.ti 7 user guide and reference* ATLAS.ti Scientific Software Development GmbH. [http://downloads.atlasti.com/docs/manual/atlasti\\_v7\\_manual\\_en.pdf?\\_ga=2.132805390.1617080549.1512772864-1300464563.1512772864](http://downloads.atlasti.com/docs/manual/atlasti_v7_manual_en.pdf?_ga=2.132805390.1617080549.1512772864-1300464563.1512772864)
- Hargreaves, A., & Dawe, R. (1990). Paths of professional development: Contrived collegiality, collaborative culture, and the case of peer coaching. *Teaching and Teacher Education*, *6*(3), 227–241.
- Hargreaves, E., & Elhawary, D. (2019). Professional development through mutually respectful relationship: Senior teachers' learning against the backdrop of hierarchical relationships. *Professional Development in Education*, *45*(1), 46–58.
- Heineke, S. F. (2013). Coaching discourse: Supporting teachers' professional learning. *Elementary School Journal*, *113*(3), 409–433.
- Helms-Lorenz, M., van de Grift, W., Canrinus, E., Maulana, R., & van Veen, K. (2018). Evaluation of the behavioral and affective outcomes of novice teachers working in professional development schools versus non-professional development schools. *Studies in Educational Evaluation*, *56*, 8–20.
- Helms-Lorenz, M., van de Grift, W., & Maulana, R. (2016). Longitudinal effects of induction on teaching skills and attrition rates of beginning teacher. *School Effectiveness and School Improvement*, *27*(2), 178–204.
- Hornsey, M. J., Oppes, T., & Svensson, A. (2002). "It's OK if we say it, but you can't": Responses to intergroup and intragroup criticism. *European Journal of Social Psychology*, *32*(3), 293–307.
- Imants, J., & van Veen, K. (2010). Teacher learning as workplace learning. In P. Peterson, E. Baker, & B. McGaw (Eds.), *International encyclopedia of education* (pp. 569–574). Elsevier.
- Kennedy, M. M. (2010). Attribution error and the quest for teacher quality. *Educational Researcher*, *39*(8), 591–598.
- Kennedy, M. M. (2016). How does professional development improve teaching? *Review of Educational Research*, *86*(4), 945–980.
- Knapp, M. S. (2003). Professional development as a policy pathway. *Review of Research in Education*, *27*(1), 109–157.
- Kuusisaari, H. (2014). Teachers at the zone of proximal development: Collaboration promoting or hindering the development process. *Teaching and Teacher Education*, *43*, 46–57.

- Leary, M., & Terry, M. (2012). Interpersonal aspects of receiving evaluative feedback. In R. M. Sutton, M. J. Hornsey, & K. M. Douglas (Eds.), *Feedback: The communication of praise, criticism, and advice* (pp. 15–41). Peter Lang.
- Linder, S. M. (2011). The facilitator's role in elementary mathematics professional development. *Mathematics Teacher Education and Development*, *13*(2), 44–66.
- Lofthouse, R. (2019). Coaching in education: A professional development process in formation. *Professional Development in Education*, *45*(1), 33–45.
- Marsh, J. A., McCombs, J. S., & Martorell, F. (2010). How instructional coaches support data-driven decision making: Policy implementation and effects in Florida middle schools. *Educational Policy*, *24*(6), 872–907.
- Maulana, R., Helms-Lorenz, M., & van de Grift, W. (2015a). Development and evaluation of a questionnaire measuring pre-service teachers' teaching behavior: A Rasch modeling approach. *School Effectiveness and School Improvement*, *26*(2), 169–194.
- Maulana, R., Helms-Lorenz, M., & van de Grift, W. (2015b). 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.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. Sage.
- Neuman, S. B., & Cunningham, L. (2009). The impact of professional development and coaching on early language and literacy instructional practices. *American Educational Research Journal*, *46*(2), 532–566.
- Neuman, S. B., & Wright, T. S. (2010). Promoting language and literacy development for early childhood educators: A mixed-methods study of coursework and coaching. *Elementary School Journal*, *111*(1), 63–86.
- Opfer, V. D., & Pedder, D. (2011). Conceptualizing teacher professional learning. *Review of Educational Research*, *81*(3), 376–407.
- Sailors, M., Hoffman, J. V., David Pearson, P., McClung, N., Shin, J., Phiri, L. M., & Saka, T. (2014). Supporting change in literacy instruction in Malawi. *Reading Research Quarterly*, *49*(2), 209–231.
- Sailors, M., & Price, L. (2015). Support for the improvement of practices through intensive coaching (SIPIC): A model of coaching for improving reading instruction and reading achievement. *Teaching and Teacher Education*, *45*, 115–127.
- Santagata, R., & Bray, W. (2016). Professional development processes that promote teacher change: The case of a video-based program focused on leveraging students' mathematical errors. *Professional Development in Education*, *42*(4), 547–568.
- Thompson, C. L., & Zeuli, J. S. (1999). The frame and the tapestry: Standards-based reform and professional development. In L. Darling-Hammond & G. Sykes (Eds.), *Teaching as the learning profession. Handbook of policy and practice* (pp. 341–375). Jossey-Bass.
- van de Grift, W., Helms-Lorenz, M., & Maulana, R. (2014). Teaching skills of student teachers: Calibration of an evaluation instrument and its value in predicting student academic engagement. *Studies in Educational Evaluation*, *43*, 150–159.
- van der Lans, R. M., van de Grift, W. J., & van Veen, K. (2018). Developing an instrument for teacher feedback: Using the Rasch model to explore teachers' development of effective teaching strategies and behaviors. *Journal of Experimental Education*, *86*(2), 247–264.
- van Driel, J. H., Meirink, J., van Veen, K., & Zwart, R. (2012). Current trends and missing links in studies on teacher professional development in science education: A review of design features and quality of research. *Studies in Science Education*, *48*(2), 129–160.
- Veenman, S., & Denessen, E. (2001). The coaching of teachers: Results of five training studies. *Educational Research and Evaluation*, *7*(4), 385–417.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Warford, M. K. (2011). The zone of proximal teacher development. *Teaching and Teacher Education*, *27*(2), 252–258.