



Universiteit
Leiden
The Netherlands

Effects of formative assessment programmes on teachers' knowledge about supporting students' reflection

Tigelaar, E.H.; Sins, P.

Citation

Tigelaar, E. H., & Sins, P. (2020). Effects of formative assessment programmes on teachers' knowledge about supporting students' reflection. *Journal Of Vocational Education And Training*, 73(3), 413-435. doi:10.1080/13636820.2020.1726992

Version: Publisher's Version

License: [Creative Commons CC BY-NC-ND 4.0 license](https://creativecommons.org/licenses/by-nc-nd/4.0/)

Downloaded from: <https://hdl.handle.net/1887/3274063>

Note: To cite this publication please use the final published version (if applicable).



Effects of formative assessment programmes on teachers' knowledge about supporting students' reflection

Dineke Tigelaar & Patrick Sins

To cite this article: Dineke Tigelaar & Patrick Sins (2021) Effects of formative assessment programmes on teachers' knowledge about supporting students' reflection, Journal of Vocational Education & Training, 73:3, 413-435, DOI: [10.1080/13636820.2020.1726992](https://doi.org/10.1080/13636820.2020.1726992)

To link to this article: <https://doi.org/10.1080/13636820.2020.1726992>



© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



[View supplementary material](#)



Published online: 16 Feb 2020.



[Submit your article to this journal](#)



Article views: 3210



[View related articles](#)



[View Crossmark data](#)



Citing articles: 1 [View citing articles](#)

ARTICLE

 OPEN ACCESS

 Check for updates

Effects of formative assessment programmes on teachers' knowledge about supporting students' reflection

Dineke Tigelaar^a and Patrick Sins^{a,b}

^aICLON, Leiden University Graduate School of Teaching, Leiden University, Leiden, The Netherlands;

^bSaxion, School of Education, Saxion University of Applied Sciences, Enschede, The Netherlands

ABSTRACT

Objective: In the past decade, several authors have advocated that formative assessment programmes have an impact on teachers' knowledge. Consequently, various requirements have been proposed in the literature for the design of these programmes. Only few studies, however, have focused on a direct comparison between programmes with respect to differences observed in their effect on teachers' knowledge. Therefore in this study we explored the impact of three formative assessment programmes on teachers' knowledge about supporting students' reflection.

Methods: Our study was carried out in the domain of vocational nursing education. Teachers were assigned to an expertise-based assessment programme, a self-assessment combined with collegial feedback programme, or a negotiated assessment programme. We scored the verbal transcriptions of teachers' responses to video vignette interviews in order to measure their knowledge in a pre- and post-test. Multilevel regression analyses were performed to investigate differences in teachers' knowledge between the three programmes on the post-test; potential moderating effects of pre-test scores, contextual and individual factors were controlled for. **Findings:** The knowledge of teachers participating in the expertise-based assessment programme was significantly higher than that of teachers participating in the self-assessment combined with collegial feedback programme. Furthermore, the findings indicate that for professional learning, not only the approach to formative assessment is an important variable, but also the extent to which (a) teachers are intrinsically motivated and (b) they experience a high degree of collegiality at their school.

ARTICLE HISTORY

Received 13 June 2019

Accepted 21 January 2020

KEYWORDS


Teacher development; assessment; vocational and educational guidance; learning in the professions; pedagogy

Introduction

This study explores the impact of three different formative assessment programmes on teachers' knowledge with respect to fostering reflection in vocational students. Formative assessment is seen as an endeavour that may offer

CONTACT Dineke Tigelaar  DTigelaar@iclon.leidenuniv.nl

This article has been republished with minor changes. These changes do not impact the academic content of the article.

 Supplemental data for this article can be accessed [here](#).

© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

effective ways to support and train teachers (Porter, Youngs, and Odden 2001; Lynch, Mannix McNamara, and Seery 2012). Empirical studies on teacher learning have mostly focused on investigating particular features of programmes that contribute to teacher professional development (Garet et al. 2001; Lusttich and Sykes 2006; Penuel et al. 2007; Sato, Wei, and Darling-Hammond 2008; Van den Bergh, Ros, and Beijaard 2014). However, very few studies have focused on a direct comparison between programmes (e.g. Heller et al. 2012).

As a content domain of our study, we focused on vocational nursing education, since developing students' reflective skills and assisting teachers in guiding their own student to develop these skills is commonly considered to be especially important (Mann, Gordon, and MacLeod 2009). Fostering nursing students' reflection means that students are supported by teachers in a conscious, dynamic process of thinking about, analysing, and learning from a clinical experience in order to respond to clinical situations with a changed perspective (Mann, Gordon, and MacLeod 2009). Reflection is perceived as a way for nurses to investigate the depth and complexity of their professional practice, understand why they do what they do, grasp the creativity of practice, explore the emotional aspects of a situation and attain a rich understanding of nurse-patient interactions (Thompson and Burns 2008). Developing reflection skills in students requires extensive assistance and scaffolding by their teachers (Asselin 2011). A vital concept in this respect is teachers' knowledge, because it is commonly acknowledged that teachers' knowledge strongly influences their teaching practice (Cochran-Smith and Zeichner 2005; Evans and Kozhevnikova 2011). Knowing more about the knowledge that teachers draw on when fostering reflection in their students and, more specifically, about how formative assessment processes have an impact on teachers' knowledge, can contribute to inspire programmes of teacher formative assessment focused on reflective pedagogies.

Below, we first conceptualise teachers' knowledge for fostering their students' reflections. Secondly, we describe our operationalisation of formative assessment as a process for supporting the development of teachers' knowledge. Finally, we describe the three formative assessment programmes that are the focus of this study.

Theoretical framework

Conceptualising teachers' knowledge about supporting reflection

The importance of teachers' knowledge was emphasised in the seminal work of Shulman (1986, 1987), who conceptualised teachers' knowledge as a particular combination of both subject knowledge and pedagogy that affects and is affected by teachers' teaching practice. Several researchers have used Shulman's work, conceptualising teacher knowledge as the integrative

aggregate of beliefs, cognitions, and knowledge that teachers draw on and generate when they encounter difficult situations or problems in their teaching practice (Verloop, Van Driel, and Meijer 2001).

In this study, we used Kreber and Cranton (1997, 2000) classification system, consisting of pedagogical knowledge, instructional knowledge, and curricular knowledge. To describe teachers' knowledge that is specific to stimulating students' reflection, we applied Kreber and Cranton's knowledge system to what is known about reflection. First, *pedagogical knowledge* refers to the knowledge required for being responsive to a particular student's level and which is aimed at diagnosing the reflection skills of a particular student (e.g. Pratt and Savoy-Levine 1998; Van de Pol, Volman, and Beishuizen 2011). Second, teachers' *instructional knowledge* is manifest in the articulation of teaching strategies aimed at stimulating cognitive activities, such as critical analysis, synthesis, and evaluation as vital skills for reflection (e.g. Atkins and Murphy 1993; Boud, Cressy, and Dochtery 2006; Oosterbaan et al. 2010). Third, the idea of *curricular knowledge* corresponds to Schoenfeld's (1998) notion of a teacher's action plan, which involves a representation of how teachers propose to achieve a specific goal with respect to the learning outcomes of their students' reflection activities. Action plans might be seen as a means towards achieving this end and are guided by teachers' rationales. The ultimate goal of reflection is to improve current practices and develop the capacity to direct one's own development (Schön 1983).

Formative assessment processes and the effect on teachers' knowledge

Formative assessment with teachers as learners is a relatively new area, since the literature on formative assessment mainly concerns research with students. Although different conceptualisations of formative assessment exist in the literature, a number of key features can be mentioned. A first key feature is that evidence of learning is used to reflect and adapt learning (Black and Wiliam 1998). This evidence is elicited and interpreted, and action is taken that results in advancement of learning (Black and Wiliam 2009). This means that formative assessment is characterised as a process (Popham 2006) with a focus on the purposeful use of assessment information by learners and their educators to adjust what they are currently doing.

Feedback is a second key feature in formative assessment processes to understand how current functioning relates to what is expected or desired and what steps can be taken in order to bridge this gap (Sadler 1989). Principles of scaffolding have been considered important in the feedback so as to provide just enough support for learners to make the step to the next level in employing a learning activity which they could not have made on their own, and to gradually decrease the amount of support and give more responsibility and control to the learner (Vygotsky 1978; May 2013). Models for formative

assessment processes have been proposed in which clarifying intentions, goals and expectations, collecting information on learners' responses to instructional stimuli, interpreting and discussing evidence of learning, and deciding upon actions and adjustments, are separate phases to be applied in a cyclic process (e.g. Ruiz-Primo and Furtak 2007; Antoniou and James 2014).

In addition, formative assessment processes should provide teachers with feedback that not only supports them in recognising strong points and areas to improve, but also challenges their assumptions, improves their understanding of their own teaching practices, and supports them in setting learning goals and choosing learning activities that help them to plan their future learning and move forward to anticipated learning benefits (Porter, Youngs, and Odden 2001). Focusing on teacher knowledge means taking into account that this knowledge is mostly tacit and difficult to articulate (Beijaard and Verloop 1996; Schön 1983; Van Driel and Berry 2010). Exchanging ideas with colleagues, seeking feedback on teaching experiences, and consulting theories of teaching in relation to practical experiences, are important activities in teachers' professional learning (Bakkenes, Vermunt, and Wubbels 2010; Meirink, Meijer, and Verloop 2007), and are considered highly supportive for the development of teachers' knowledge (Van Driel and Berry 2010).

Reasoning from what we know about formative assessment, teacher learning and teacher professional development, programmes of teacher formative assessment may differ on two dimensions: (a) practice- versus research-based or (b) teacher-directed versus trainer-directed.

The first dimension, i.e. practice-based versus research-based, implies that either teachers' own practical concerns and practical knowledge are at the centre of attention, or research-based knowledge has a central focus. Although evidence-based theoretical notions derived from research may provide teachers with relevant content to the development of their knowledge, teachers hardly make use of theoretical notions in their own practice (Janssen, Westbroek, and Doyle 2015). Taking teachers' individual practice and teachers' personal concerns as a point of departure may improve teachers' sense of ownership and relevance for their own practice (Uhlenbeck, Verloop, and Beijaard 2002). In practice-based programmes of teacher formative assessment, peer teachers may be involved as feedback agents so as to ensure that assessment information is interpreted and discussed in a way that matches with teachers' concerns. Teachers themselves can also be feedback agents, e.g. through self-assessment. Self-assessment is seen as a promising approach to teacher professional development (Airasian et al. 1995; Ross and Bruce 2007; Van Diggelen, den Brok, and Beijaard 2013), since it is expected to stimulate self-monitoring and learning from experiences (Boud, Cohen, and Sampson 1999). Receiving feedback from colleagues through peer-assessment can complement teachers' ability to consider their own teaching practice critically (Ross and Bruce 2007; Lynch, Mannix McNamara, and Seery 2012) and critiquing each

other's work based on classroom observations is known to be valuable for making teacher assessment beneficial for teacher professional learning (Lusttack and Sykes 2006; Sato, Wei, and Darling-Hammond 2008; Lustick 2011).

In research-based programmes of teacher formative assessment, on the other hand, feedback agents with domain expertise may be involved, such as an experienced colleague or a trainer with knowledge in the field to ensure that feedback provides teachers with enriched insights or state-of the art knowledge. Guidance by feedback agents with domain knowledge is considered beneficial for supporting expertise development (Degner and Gruber 2011). Professional development is also known to be more effective if it is informed by research-based knowledge on a particular content area (Desimone 2009; Garet et al. 2001; Penuel et al. 2007; Van Driel and Berry 2010). Furthermore, implementing state-of the art knowledge can improve the quality of observations, evaluative judgements and feedback (e.g. Heller, Sheingold, and Myford 1998; Hattie and Timperley 2007). Although empirical evidence is scarce (Van Diggelen, den Brok, and Beijaard 2013), combining self-assessment and peer-assessment and providing teachers with research-based instruments and guidelines seems promising for finding an optimal balance between practice-based and research-based formative teacher assessment.

The second dimension, i.e. teacher-directed versus more trainer-directed, implies that responsibility and control is either mostly with the teachers or mostly with the trainer. Giving teachers a considerable amount of control is more in line with putting emphasis on using assessment information for setting personal learning goals, co-constructing success criteria and developing self-regulated learning skills for monitoring learning processes (also referred to as 'assessment as learning') (Clark 2012). Trainer-direction is more compatible with using assessment information for diagnosing knowledge and skills and possible gaps between current functioning in relation to certain standards (Clark 2012). In line with literature emphasising that learners should actively and deliberately take responsibility over their own learning through formative assessment so as to develop their self-regulation skills (Stiggins and Chappuis 2005; Clark 2012), a certain amount of guidance in formative assessment of teaching seems useful as a way to scaffold teacher learning and/or point them towards the learning needs that they might not yet be conscious of (Desimone 2009, cf. May 2013). Including opportunities for negotiation in formative teacher assessment programmes can motivate teachers to be actively involved and to focus more on their personal learning goals and concerns, although the feedback agent may need to challenge teachers to take these opportunities (Anderson, Boud, and Sampson 1996; Gosling 2000; De Eça and Torres 2005). As such, negotiated formative assessment may be valuable for finding an optimal balance between trainer-guidance and teacher guidance in programmes of teacher formative assessment.

The current study

In this study, we explored the differential effect of three different programmes of teacher formative assessment that were designed for supporting teachers in vocational nursing education in fostering their students to reflect.

Three programmes of teacher formative assessment

The following three programmes were studied (see [Table 1](#)):

- (1) Feedback on reflection involving an expertise-based formative assessment, in which teachers received training and feedback from a certified trainer, both in individual and group sessions.
- (2) A formative assessment procedure for Self Assessment combined with Feedback from Colleagues, in which fixed pairs of teachers assessed themselves and a colleague and provided each other with feedback.
- (3) A formative assessment procedure for Negotiation-based Assessment, in which fixed pairs of a teacher (i.e. assessee) and an experienced colleague (i.e. assessor) negotiated about the evidence of learning, the feedback and suggestions provided by the assessor as well as about teachers' learning objectives, and their follow up actions.

With regards to the two dimensions (i.e. trainer-directed versus teacher-directed and research-based versus practice-based) the design of the expertise-based assessment programme was the most trainer-directed and research-based, the design of self-assessment and collegial feedback programme was the most teacher-directed, while situated in the midst of the research-based versus practice-based dimension, and the design of the negotiated assessment programme was positioned in the midst of the trainer-directed versus teacher-directed dimension, and the most practice-based (see also [Table 1](#)). Below, we describe in more detail how the key features of formative assessment processes were operationalised in the three programmes.

Expertise-based assessment

In the expertise-based programme, the feedback agent was a certified trainer with experience in nursing who provided teachers with feedback based on research-based information grounded in both the available research literature and knowledge from experts in the field of reflection (Dekker-Groen, Van der Schaaf, and Stokking 2013). During group meetings with the three participating teachers at each school, on the basis of criteria and standards derived from the literature the trainer explicitly reported on relevant research-based knowledge with respect to, e.g. asking questions aimed at stimulating students' reflection and engaging in entertaining reflective dialogues. During individual feedback meetings, each teacher was

Table 1. The three formative assessment programmes.

	Expertise-based Assessment	Self-Assessment & Collegial Feedback	Negotiated Assessment
Feedback setting	<p>A certified trainer with experience in nursing (i.e. the expert) reports during group meetings on relevant research-based knowledge with respect to stimulating students' reflection, based on criteria and standards derived from literature.</p> <p>During individual feedback meetings, teachers reflect on a video recording of a reflection conversation with one of their students, and the trainer asks questions and refers to theoretical notions discussed in the group meetings.</p> <p>During individual feedback meetings, the trainer defines the relevant goals and learning activities for each teacher.</p>	<p>Teachers with similar experience work in fixed pairs and provide each other with peer feedback during collegial feedback meetings. Teachers also assess themselves.</p> <p>Feedback is provided based on observations and assessment of a video recording of a reflection conversation with one of their students. Teachers use a tool based on a conceptual model about stimulating students' reflection skills to assess themselves and their peers. A structured conversation instrument is used for providing feedback.</p> <p>At the end of each collegial feedback session, teachers set their own learning goals and after each session the teachers write a reflective report, in which they integrate their learning experiences and formulated areas for improvement.</p> <p>Videos of reflection conversations with students (required).</p>	<p>During negotiated assessment meetings, a teacher and a more experienced colleague from the same school negotiate their interpretations of a video recording of a reflection conversation from the teacher with a particular student. Also the feedback and suggestions are negotiated.</p> <p>A framework for teaching competences with regards to stimulating students' reflection skills is used for the interpretations and negotiations which is based on the input of various stakeholders in the field, including teachers themselves.</p> <p>At the end of each negotiated assessment meeting, the assessor and the teacher negotiate the learning goals, activities to be carried out and further evidence of learning to be collected.</p> <p>Evidence to be collected is negotiated with the experienced colleague.</p>
Evidence	<p>Videos of reflection conversations with students (required).</p>	<p>Videos of reflection conversations with students (required).</p>	<p>Evidence to be collected is negotiated with the experienced colleague.</p>

stimulated to reflect on a video recording of a reflection conversation he or she had conducted with a particular student, and the trainer asked questions and made references to theoretical notions discussed in the group meetings. During these individual feedback meetings, the trainer also defined the relevant goals and learning activities for each teacher with regard to fostering students' reflection in order to stimulate the teachers to take action based on the feedback provided.

Self-assessment combined with feedback from colleagues

In the programme for Self-assessment combined with Feedback from Colleagues, colleague teachers from the same school with similar experience were the feedback agents. These teachers assessed themselves and, in fixed pairs, they also provided each other with peer feedback during collegial feedback meetings (Van Diggelen, den Brok, and Beijaard 2013). During each collegial feedback meeting, the teachers provided each other with feedback based on their observations and assessment of a video of a reflection conversation each teacher had conducted with a particular student. To provide guidance and structure, teachers were provided a tool to assess themselves and their peers, which was based on a conceptual model about stimulating students' reflection skills. Teachers used this instrument and the conceptual framework to assess themselves, and their colleague teacher. Research-based knowledge was also brought in by handing the teachers a structured conversation instrument that was meant to assist the teachers in providing good quality feedback which they were required to use during collegial feedback discussions. The self-assessments and peer feedback during collegial feedback meetings were carried out by the teachers without the presence of a trainer or a feedback agent with specific expertise in the field. At the end of each collegial feedback session, teachers set their own learning goals with respect to stimulating students' reflection in order to stimulate the teachers to take action based on the feedback provided. In addition, after each session the teachers wrote a reflective report, in which they integrated their learning experiences and formulated areas for improvement.

Negotiated assessment programme

The programme for negotiated assessment was carried out in fixed pairs, i.e. a teacher and a more experienced colleague within the same school (the latter being the feedback agent) (Verberg, Tigelaar, and Verloop 2013). During each negotiated assessment meeting, the teacher and the more experienced colleague were stimulated to negotiate their interpretations of a video recording of a reflection conversation the teacher had conducted with a particular student. While negotiating, the teachers were stimulated to take their own practical concerns with regards to stimulating students' reflections as central focus, and to negotiate with their experienced colleague not only the interpretations with regards to the evidence of learning, but also the feedback and suggestions provided. At the end of each negotiated assessment meeting, the learning

goals, the activities to be carried out in order to achieve these goals and further evidence of learning to be collected in order to demonstrate development, were negotiated between the assessor and the teacher. As a stimulus for interpreting the evidence of learning and carrying out the negotiations, teachers were provided with a framework for teaching competences with regards to stimulating students' reflection skills which was based on the input of various stakeholders in the field, including teachers themselves.

Research focus

As mentioned above, the research is focused on exploring the impact of three formative assessment programmes on teachers' knowledge with regards to stimulating vocational nursing students' reflection skills. To gain information on potentially contextual and individual factors moderating the effect of the programmes on the development of teachers' knowledge, we also considered, respectively, conditions for professional learning at school and situational motivation. Firstly, conditions for learning at school is a construct focused on the perceived features of the institutional context of a teacher, consisting of material, cultural, and/or social structures in the organisation that are focused on facilitating formal and informal learning. Teachers' perceptions of school conditions such as supportive leadership and professional dialogue for example, have been found to be related to teacher professional learning and the goals that teacher set for their learning (e.g. Louws et al. 2017). Other research (e.g. Runhaar and Sanders 2013) has shown that the extent to which teachers are provided with feedback, is dependent on school leaders' views on teacher evaluation in relation to professional development. Furthermore, Lillejord and Børte (2019) found that when school leaders see teacher evaluation as a possibility to promote teacher professional learning, teacher evaluation may be better facilitated and feedback given more often.

Secondly, we included situational motivation as possible moderator. Situational motivation refers to the motivation that relates to the experience of being involved in a particular activity at a particular moment in time (Vallerand 1997). The construct originates from Deci and Ryan (1985), Ryan and Deci (2000) self-determination theory, which states that in order to contribute to an individual's situational motivation, basic psychological needs concerning competence, autonomy and relatedness need to be attended to in the social context. This means that social contexts that satisfy these basic needs contribute to fostering self-determined types of motivation (i.e. intrinsic motivation and identified regulation), whereas contexts that constrain these needs may result in non-self-determined motivation types (i.e. external regulation and amotivation). Research has shown that teachers who are intrinsically motivated to work on their professional

learning, are also more engaged in professional development activities (e.g. Jansen In de Wal et al. 2014).

We formulated the following research question to investigate the impact of the three programmes on the development of teachers' knowledge: *What are the differential effects of three formative assessment programmes on teachers' knowledge about stimulating vocational nursing students' reflection skills?*

Methods

Participants and context

The participants were teachers in Dutch senior secondary vocational education for undergraduate nursing students (level IV of the International Standard Classification of Education, UNESCO).

In this research, we studied the differential effects of the three formative assessment programmes on the knowledge of thirty-seven teachers in secondary vocational education for nursing: 12 teachers from four different schools in the expertise-based assessment programme, 12 teachers from three schools in the negotiated assessment programme and 13 teachers from two schools in the self-assessment combined with collegial feedback programme; 11 teachers were male, 26 were female. Their average age was 49 (SD = 5.74) and their average teaching experience was 12 years (SD = 8.33). Participants were informed about the aims of our research and participated on a voluntary basis. They were informed that their responses were not being evaluated in connection with student reflective ability and that their information was kept confidential.

Procedures and instruments

Formative assessment programmes

The three programmes were executed at different schools, and the arrangement, timing, and duration of each round were attuned to the time schedules of the participating schools. Each session took about 1 to 1,5 hour. In total, in each programme three consecutive rounds were organised. On the whole each programme lasted for about two years, and time between each round was approximately one year. The teachers did not have other assessments with regards to their proficiency in stimulating students' reflection skills during that time. In between the sessions, teachers had the opportunity to experiment with suggestions based on the feedback provided.

Video vignettes for measuring teachers' knowledge

To explore the development of teachers' knowledge, we measured teachers' knowledge using six video vignette interviews, which were used in a pre-test

Table 2. Example of a vignette.

Returning from one of her first internships, Susan says that she's noticed that many staff members do not follow the hygiene measures taught at school. Most people, for instance, work with watches on their wrists and with long polished nails, while at school she was taught that this is unhygienic. Susan is confused and indignant. Susan is a first-year student who normally obtains good results. She says: 'What they tell us at school is so outdated!' (video) Subsequent reaction from Susan: 'Well, but nobody says it's wrong!' (video) Imagine Susan subsequently says: 'Now I'm confused, at school I learnt this and when I go on internship they do it differently. What should I do now?' Imagine Susan subsequently says: 'What can be wrong with long nails? I wash my hands five times a day.'

and post-test setting. The vignettes and the interview protocol were designed in an earlier study (Tigelaar, Sins, and van Driel 2017) to present teachers with multiple authentic situations that were prototypical in their teaching context and to elicit teaching interventions (e.g. questions they would ask and feedback they would give for supporting students' reflection in that typical situations) and teachers' rationales behind these interventions (more details on the development of the video vignettes and the interview protocol can be found in Tigelaar, Sins, and van Driel 2017; see also Table 2 for an example of a video vignette, and Table S1 in the supplementary materials accompanying the online article, for the complete overview of the vignettes).

Each vignette consists of contextual information describing a particular critical situation, and four typical student utterances. After viewing each student utterance, teachers were prompted using a set of structured questions (i.e. 'How would you react to this situation?'; 'Why?'; 'What do you intend to promote in terms of student learning with these interventions?'; and 'How would you handle this situation?'). Before each of the three programmes for formative assessment commenced, we administered the video vignette interviews for measuring teachers' knowledge (i.e. pre-test). After the three programmes for formative assessment had finished, we repeated the procedure with the video vignettes with the teachers (i.e. post-test). Each interview took about one hour.

Conditions for professional learning at school

Based on a synthesis of the literature focused on assessing conditions for learning at school (Ellström 2001; Kyndt, Dochy, and Nijs 2009; Geijsel et al. 2001; Baert, De Witte, and Sterck 2000), we derived the following four scales consisting of, respectively, 6, 13, 3, 6, and 8 items: (1) feedback and opportunities for exchanging information and knowledge (e.g. 'At our school I have the opportunity to collaborate with colleagues'); (2) characteristics of the organisation (e.g. 'At our school teachers are stimulated to experiment with new didactic methods'); (3) coaching (e.g. 'At our school I am guided in my professional development during my performance review'); and (4) available tools and resources (e.g. 'At our school I have the opportunity to use the internet and ICT for my own professional development'). Teachers were asked to respond on a five-point Likert scale, varying from 'totally

disagree' to 'totally agree'. Coefficient alphas for the four scales varied between .71 and .80.

Situational motivation

Guay, Vallerand, and Blanchard (2000) situational motivation scale (SIMS) was designed to assess the constructs of intrinsic motivation (e.g. 'Because I think that this activity is interesting'), identified regulation (e.g. 'Because I am doing it for my own good'), external regulation ('Because I am supposed to do it'), and amotivation (e.g. 'There may be good reasons to do this activity, but personally I don't see any'). Each scale consists of four items and prompts respondents to indicate the rationale for being engaged in a particular activity by responding on a five-point Likert scale, varying from 'does not correspond at all' to 'corresponds exactly'. Guay, Vallerand, and Blanchard (2000) original situational motivation scale used a seven-point Likert scale, which we adapted to a five-point Likert scale.

Coefficient alphas for the four scales varied between .70 and .93.

Teachers were asked to complete the questionnaires aimed at determining conditions for professional learning and situational motivation directly after the final round of each formative assessment programme.

Analysis

Analysis of the video vignette data

To analyse the development of teachers' knowledge, verbal protocols were obtained based on the transcribed voice recordings of teachers' responses to the vignettes at both pre- and post-test. We analysed the interview transcripts from the pre- and post-test using five criteria of teacher knowledge for supporting reflection. For each of the five criteria, four standards were defined (see Table 3).

To substantiate our holistic scoring of the transcripts, we used the relative frequencies of codes assigned to segments identified in the transcripts. These codes were derived from a protocol analysis scheme developed by Tigelaar, Sins, and van Driel (2017). Interrater reliability of this scheme was considered acceptable (Cohen's kappa = .75). For this procedure, the interview transcripts were segmented into meaningful units composed of coherent continuous talk on a single topic or theme (cf. Chi 1997). Segments were scored by assigning a code for curricular, pedagogical or instructional knowledge and a code for reference to reflection objects. For the first criterion (degree of contingent coaching), we analysed the extent to which teachers were responsive to a particular student's level (i.e. pedagogical knowledge; Pratt and Savoy-Levine 1998; Van de Pol, Volman, and Beishuizen 2011). The second criterion (degree of stimulating students' reflection through instructional strategies) included a measure of teachers' instructional knowledge and involved assessing the extent to which teachers aimed at fostering students' higher-order

Table 3. Explanation of standards for the five criteria for scoring teachers' knowledge.

1	2	3	4
<p>Criterion 1: Degree of contingent coaching The teacher provides no sign of constructing diagnoses of students' reflective skills</p>	<p>The teacher mentions diagnoses, but provides no sign of checking them. In addition, these diagnoses are based on a singular assumption concerning students' reflective skills</p>	<p>The teacher provides several diagnoses based on a number of assumptions concerning students' reflective skills</p>	<p>The teacher reasons frequently from confirmatory as well as differential diagnoses based on a heterogeneous set of assumptions concerning students' reflective skills</p>
<p>Criterion 2: Degree of stimulating students' reflection through instructional strategies The teacher is unilaterally focused on transferring knowledge to students</p>	<p>The teacher mostly articulates instructional strategies focusing on stimulating students' descriptive reflection activities</p>	<p>The teacher regularly articulates instructional strategies focusing on stimulating students' descriptive and dialogical reflection</p>	<p>The teacher frequently and evenly articulates instructional strategies focusing on stimulating students' descriptive, dialogical, and critical reflection</p>
<p>Criterion 3: Degree of attentiveness to reflection in a safe environment The teacher barely attends to stimulating students' reflection in a safe environment</p>	<p>The teacher provides some affordances for stimulating students' reflection in a safe environment, which are barely based on assumptions concerning their reflective skills</p>	<p>The teacher provides several affordances for stimulating students' reflection in a safe environment, which are based on assumptions concerning their reflective skills</p>	<p>The teacher is highly attentive to stimulating students' reflection in a safe environment, based on assumptions concerning their reflective skills</p>
<p>Criterion 4: Degree of focus on achieving students' higher-order reflection goals The teacher barely attends to stimulating students' higher-order reflection goals</p>	<p>The teacher provides some affordances for stimulating students' higher-order reflection goals</p>	<p>The teacher provides several affordances for stimulating students' higher-order reflection goals</p>	<p>The teacher is highly attentive to stimulating students' higher-order reflection goals</p>
<p>Criterion 5: Degree of focus on students' reflection on professional behaviour The teacher barely attends to stimulating reflection on professional behaviour</p>	<p>The teacher provides some affordances for stimulating reflection on professional behaviour and instead focuses on stimulating students' reflection on observed behaviour</p>	<p>The teacher provides several affordances for stimulating reflection on professional behaviour</p>	<p>The teacher is highly attentive to stimulating reflection on professional behaviour and knowledge</p>

reflection through their interventions (Atkins and Murphy 1993; Boud, Cressey, and Dochtery 2006; Oosterbaan et al. 2010). The third criterion (degree of attentiveness to reflection in a safe environment) comprised a combination of teachers' pedagogical and instructional knowledge, focusing on determining the degree to which teachers aimed at maintaining a safe environment contingent with students' needs (Asselin 2011; Thompson and Burns 2008). The fourth criterion (degree of focus on achieving students' higher-order reflection goals) reflects a measure of teachers' curricular knowledge, determining the extent to which teachers intend to achieve higher-order reflection goals (Mann, Gordon, and MacLeod 2009). Finally, the fifth criterion (degree of focus on students' reflection on professional behaviour in nursing) was aimed at assessing the object of reflection and, more specifically, the degree of teachers' focus on stimulating students' reflection on their professional identity (Mann, Gordon, and MacLeod 2009). Interrater reliability for the scored standards was determined by comparing the ratings of two independent judges on all five criteria ($n = 15$; Cohen's kappa = .61; percentage agreement = 73%). These interrater reliability coefficients can be considered moderate based on Fleiss (1971) and McHugh (2012).

Scores on the items obtained from the questionnaires for measuring conditions for professional learning at school and situational motivation were aggregated into mean scores for each of their respective scales.

Analyses of differential effects on teachers' knowledge

Multilevel regression analyses were performed to investigate differences in the effects of the three programmes on teachers' knowledge. We controlled for potential moderating effects of contextual and individual factors. We performed univariate multilevel regression analyses with the post-test score of teachers' knowledge as dependent variable. The model had two levels: teacher and school. However, the school level had to be excluded, since variance on this level was found to be zero. To investigate differences in the impact of the three programmes on teacher knowledge, we controlled for teachers' pre-test scores.

We also intended to control for possible effects of gender and teaching experience on the development of teachers' knowledge; however, the large number of missing values precluded the use of teaching experience in the analyses. We compared five models. The first model included only the random variables and the pre-test of the knowledge measurement (first model). This model was compared with the null model, which only contained the intercept. The second included the control variable for gender. The third and fourth model added the variables of conditions for professional learning at school and situational motivation, respectively. The fifth model included the dummy variables for programme of formative assessment (i.e. expertise-based assessment, self-assessment combined with collegial feedback and negotiated assessment programme).

Table 4. Averages and standard deviations for the (a) pre-and post-test of the measurement of teachers' knowledge, (b) four scales of the conditions for professional learning at school questionnaire, and (c) four scales of the situational motivation questionnaire.

	Self-assessment											
	combined with collegial feedback (n = 13)				Expertise-based assessment (n = 12)				Negotiated assessment (n = 12)			
	pre-test		post-test		pre-test		post-test		pre-test		post-test	
	M	sd	M	sd	M	sd	M	sd	M	sd	M	Sd
Teacher knowledge	13.69	2.25	13.46	1.13	13.42	2.15	14.42	1.00	13.33	2.15	13.58	1.73
<i>Conditions for professional learning at school</i>												
Feedback and opportunities for exchange			2.71	.59			3.13	.45			3.32	.37
Characteristics of the organisation			3.14	.35			3.37	.26			3.34	.62
Coaching			3.19	.74			3.57	.62			3.53	.81
Available tools			3.41	.63			3.43	.53			2.98	.70
<i>Situational motivation</i>												
Intrinsic motivation			3.29	.82			3.06	.57			3.15	.88
Intrinsic regulation			3.19	.87			3.33	.47			3.67	.79
External regulation			2.98	.83			2.46	.74			2.61	1.08
Amotivation			2.37	1.14			2.29	.55			1.96	1.24

MLWin (version 2.23) was used to analyse the data. To determine whether adding the covariates significantly improved the model fit, Chi-square-tests were performed. The estimation method used in all the models is IGLS (maximum likelihood estimation).

Results

Table 4 presents descriptive data of teachers' knowledge measurements, the conditions for professional learning at school questionnaire, and the situational motivation questionnaire. Table 4 shows that teachers in the self-assessment combined with collegial feedback programme showed a modest decrease in scores on the knowledge measure. Paired *t*-tests showed that this decline was not significant. Knowledge scores of teachers in both the expertise-based assessment programme and the negotiated assessment programme showed a non-significant increase.

We checked whether teachers' pre-test scores differed per programme by conducting a univariate analysis with the first measurement as dependent variable. The model originally had two levels: teacher and school. However, the school level was excluded, since variance on this level was absent. We compared three models. The first model included only the random variables (first model). The second included the control variable for gender. In the third, we added dummy variables for the formative assessment programme. Addition of the dummy variables in Model 3 did not significantly improve the model fit. In addition, none of the regression weights was significant, indicating that there were no significant differences in teacher knowledge between programmes on the pre-test.



Table 5. Effects of the type of programme for formative assessment, conditions for professional learning at school, and situational motivation on the development of teachers' knowledge.

	Model 1	Model 2	Model 3	Model 4	Model 5
	Development of teachers' knowledge	Development of teachers' knowledge	Development of teachers' knowledge	Development of teachers' knowledge	Development of teachers' knowledge
Intercept	11.051 (1.347)	11.213 (1.434)	10.437 (2.115)	7.228 (3.099)	7.786 (2.910)
Pre-Test teacher professional development	0.205 (0.099)	0.200 (0.099)	0.209 (0.091)	0.255 (0.098)	0.248 (0.090)
Gender (reference = male)		-0.149 (0.457)	-0.322 (0.420)	-0.386 (0.409)	-0.426 (0.376)
<i>Conditions for professional learning at school</i>					
Feedback and opportunities for exchange			-0.526 (0.432)	-0.330 (0.428)	-0.578 (0.452)
Characteristics of the organisation			-0.212 (0.539)	-0.073 (0.535)	-0.268 (0.508)
Coaching			0.958 (0.370)	1.161 (0.445)	1.168 (0.408)
Available tools			-0.061 (0.344)	-0.407 (0.412)	-0.419 (0.420)
<i>Situational motivation</i>					
Intrinsic motivation				0.528 (0.339)	0.625 (0.312)
Identified regulation				-0.172 (0.355)	-0.210 (0.342)
External regulation				-0.087 (0.251)	0.076 (0.241)
Amotivation				0.526 (0.328)	0.441 (0.307)
<i>Approach</i> (reference = self-assessment combined with collegial feedback)					
Expertise-based assessment					1.193 (0.463)
Negotiated assessment					0.493 (0.552)
<i>Variance (residual)</i>					
Teacher	1.590 (0.370)	1.586 (0.369)	1.307 (0.304)	1.181 (0.275)	0.988 (0.230)
Percentage explained variance (R-squared)	10.4%	0.3%	17.6%	9.6%	16.3%
(vs. previous model)					
Fit	122.166	122.060	114.893	111.151	104.572
Improvement	4.066	0.106	7.167	3.742	6.579
Df and p-value	df = 1, p < .05	df = 1, p = .74	df = 4, p = .13	df = 4, p = .44	df = 2, p < .05

Bold is significant for $\alpha = .05$. Standard errors are reported in parentheses.

Table 5 shows the results of the univariate models to investigate the effects of the three programmes for formative assessment on teachers' knowledge on the post-test. As mentioned in the methods section, in the models presented, the pre-test measure was included as control variable.

We found a significant difference in between teachers in the expertise-based assessment programme and teachers who had participated in the self-assessment combined with feedback from colleagues programme. Teachers in the former programme scored significantly higher than teachers in the latter approach (see Table 5). The expertise-based assessment programme also showed a higher score on teachers' knowledge than did the negotiated assessment programme. However, this latter difference was not significant. In addition, we found that the amount of perceived coaching was a condition for professional learning at school which was significantly and positively related to the teachers' knowledge on the post-test. For the other conditions for professional learning at school, i.e. feedback and opportunities for exchanging information and knowledge, characteristics of the organisation, and available tools and resources we found a negative, though non-significant relation with teachers' knowledge. Teachers' intrinsic motivation was found to be significantly and positively associated with their knowledge. Identified regulation was found to be negatively related, however, this result was not significant. For two other motivation constructs, i.e. external regulation, and amotivation we found a positive, but also non-significant relation with teachers' knowledge. Finally, scores obtained on the pre-test of the teacher knowledge measure were significantly positively related to scores on the post-test.

Discussion

The results of this study show that the knowledge of teachers participating in the expertise-based assessment programme was significantly higher than that of teachers participating in the self-assessment combined with collegial feedback programme.

An explanation can be found in the expertise of the feedback agent involved and the guidance provided by this agent during the formative assessment processes. In the expertise-based assessment programme, the feedback agent was a certified trainer whereas in the self-assessment combined with collegial feedback programme, teachers with similar expertise were the feedback agents. Both in the expertise-based programme and the self-assessment and collegial feedback programme, research-based knowledge was important, however, in the expertise-based assessment programme, the trainer was in control of providing feedback based on theoretical notions with regards to stimulating students' reflection skills throughout the whole process. By contrast, although teachers in the self-assessment and collegial feedback programme did receive a conceptual model based on theoretical notions with regards to stimulating students' reflection skills as well as and instruments based on this model, when being engaged with the

formative assessment processes, teachers worked with these materials by *themselves*. These results suggest that trainer-directed, theory-based formative assessment provides more effective stimuli for supporting teacher professional learning when the trainer has considerable involvement throughout the assessment programme. This explanation is supported by literature that stresses the importance of feedback from agents with domain expertise as particularly valuable for supporting learning (Degner and Gruber 2011). In addition, this explanation corresponds with literature indicating that a certain amount of guidance in feedback is valuable to point teachers towards the steps they cannot take by themselves (cf. May 2013). A high degree of guidance and feedback provided by agents with domain expertise may be essential in particular in a content area that is relatively new to teachers, such as the domain of the current study (i.e. fostering reflection in students). Our finding that post-test teachers' knowledge scores of teachers in the negotiated assessment programme did *not* differ significantly from the scores of teachers in the expertise-based programme may imply that, when a formative assessment programme is mainly practice-based (i.e. the negotiated assessment programme), having a feedback agent with more expertise remains important, but control may be shared more between the feedback agents and the teachers.

Our findings showed that teachers' perceptions of the amount of perceived coaching was significantly and positively related to teachers' knowledge scores on the post-test. This means that if teachers experienced a high degree of coaching with regards to their own professional learning, for example during performance reviews, they were more likely to score higher on their knowledge related to stimulating students' reflection. This finding corroborates with the findings of other studies, which have indicated that teachers' perceptions of school conditions such as supportive leadership and professional dialogue for example, have been found to be related to their teacher professional learning (Louws et al. 2017).

In addition, teachers' intrinsic motivation was significantly and positively related to their knowledge. This means that when teachers held the opinion that participating in the formative assessment programme was interesting or valuable for them, they were more likely to score higher on their knowledge related to stimulating students' reflection. Other research has found that teachers who are intrinsically motivated to work on their professional learning, are also more engaged in professional development activities (e.g. Jansen In de Wal et al. 2014). These findings indicate that for professional learning, not only the approach to formative assessment is an important variable, but also the extent to which (a) teachers are intrinsically motivated and (b) they experience a high degree of collegiality at their school.

Limitations and suggestions for further research

A limitation is that the results could not be compared with those of a control group that did not participate in a programme for knowledge development

regarding stimulating students' reflection skills. Given that the sample size of our exploratory study was small it is possible that the differences found in this study would be larger with a larger N. Another limitation is that the percentage agreement between two independent judges was moderate, though acceptable (McHugh 2012). In further research, in order to provide additional qualitative data to explain our findings, the criteria could be complemented with concrete examples of teachers' knowledge articulations in relation to the vignettes presented in the pre- and post-test.

Furthermore, our findings do not provide results for differential effects of the assessment programmes on (a) teachers' actual behaviour in the classroom or (b) student results. Future research might focus on an in-depth investigation of professional learning activities carried out by the teachers, as well as a fine-grained analysis of interview data on how the teachers perceive the formative assessment programmes to contribute to their knowledge with regard to fostering their students to reflect.

We aspire that our exploration will be followed by many discussions on and more papers describing and evaluating different formative assessment programmes in relation to teacher professional learning in a wide variety of professional learning programmes.

Acknowledgments

We would like to thank all teachers and their schools for participating in the current study.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was supported by The Netherlands Initiative for Education Research from The Netherlands Organization for Scientific Research (NWO/PROO), under Grant number 411-06-314. The sponsor did not have any role in study design, in the data collection, analysis, and interpretation of data, or in the writing of the report, or in the decision to submit the paper for publication.

References

- Airasian, P., A. Gullickson, L. Hahn, and D. Farland. 1995. *Teacher Self-evaluation: The Literature in Perspective*. Kalamazoo: Center for Research of Educational Accountability and Teacher Evaluation.
- Anderson, G., D. Boud, and J. Sampson. 1996. *Learning Contracts*. London and New York: Routledge Farmer.

- Antoniou, P., and M. James. 2014. "Exploring Formative Assessment in Primary School Classrooms: Developing a Framework of Actions and Strategies." *Educational Assessment, Evaluation and Accountability* 26 (2): 153–176. doi:10.1007/s11092-013-9188-4.
- Asselin, M. E. 2011. "Narrative Reflection: A Method for Learning through Practice." *Journal for Nurses in Staff Development* 27 (1): 2–6. doi:10.1097/NND.0b013e3181b1ba1a.
- Atkins, S., and K. Murphy. 1993. "Reflection: A Review of the Literature." *Journal of Advanced Nursing* 18 (8): 1188–1192. doi:10.1046/j.1365-2648.1993.18081188.x.
- Baert, H., K. De Witte, and K. G. Sterck. 2000. *Vorming, Training En Opleiding. Handboek Voor Een Kwaliteitsvol VTO-beleid in Welzijnsvoorzieningen [Instruction, Training and Education (ITE). Handbook for a High Quality ITE-policy in Welfare Services]*. Leuven: Garant.
- Bakkenes, I., J. D. Vermunt, and T. Wubbels. 2010. "Teacher Learning in the Context of Educational Innovation: Learning Activities and Learning Outcomes of Experienced Teachers." *Learning and Instruction* 20 (6): 533–548. doi:10.1016/j.learninstruc.2009.09.001.
- Beijaard, D., and N. Verloop. 1996. "Assessing Teachers' Practical Knowledge." *Studies in Educational Evaluation* 22 (3): 275–286. doi:10.1016/0191-491X(96)00016-8.
- Black, P., and D. Wiliam. 1998. "Assessment and Classroom Learning." *Assessment in Education* 5 (1): 7–71.
- Black, P., and D. Wiliam. 2009. "Developing the Theory of Formative Assessment." *Educational Assessment, Evaluation and Accountability* 21 (1): 5–31. doi:10.1007/s11092-008-9068-5.
- Boud, D., P. Cressy, and P. Dochtery. 2006. "Setting the Scene for Productive Reflection." In *Productive Reflection at Work*, edited by D. Boud, P. Cressy, and P. Dochtery, 3–10. London: Routledge.
- Boud, D., R. Cohen, and J. Sampson. 1999. "Peer Learning and Assessment." *Assessment and Evaluation in Higher Education* 24 (4): 413–426. doi:10.1080/0260293990240405.
- Chi, M. T. 1997. "Quantifying Qualitative Analyses of Verbal Data: A Practical Guide." *The Journal of the Learning Sciences* 6 (3): 271–315. doi:10.1207/s15327809jls0603_1.
- Clark, I. 2012. "Formative Assessment: Assessment Is for Self-regulated Learning." *Educational Psychology Review* 24 (2): 205–249. doi:10.1007/s10648-011-9191-6.
- Cochran-Smith, M., and K. M. Zeichner. 2005. *Studying Teacher Education. The Report of the AERA Panel on Research and Teacher Education*. Washington, DC: AERA and Lawrence Erlbaum Associates.
- De Eça, M., and T. Torres. 2005. "Using Portfolios for External Assessment: An Experiment in Portugal." *International Journal of Art & Design Education* 24 (2): 209–218. doi:10.1111/j.1476-8070.2005.00441.x.
- Deci, E. L., and R. M. Ryan. 1985. *Intrinsic Motivation and Self-Determination in Human Behavior*. New York: Plenum.
- Degner, S., and H. Gruber. 2011. "Persons in the Shadow: How Guidance Works in the Acquisition of Expertise." In *The Politics of Empathy. New Interdisciplinary Perspectives on an Ancient Phenomenon*, edited by B. Weber, E. Marsal, and T. Dobashi, 103–116. Münster: Lit.
- Dekker-Groen, A. M., M. F. Van der Schaaf, and K. M. Stokking. 2013. "A Teacher Competence Development Programme for Supporting Students' Reflection Skills." *Teachers and Teaching: Theory and Practice* 19 (2): 150–171. doi:10.1080/13540602.2013.741837.
- Desimone, L. M. 2009. "Improving Impact Studies of Teachers' Professional Development: Towards Better Conceptualizations and Measures." *Educational Researcher* 38 (3): 181–199. doi:10.3102/0013189X08331140.
- Ellström, P. E. 2001. "Integrating Learning and Work: Conceptual Issues and Critical Conditions." *Human Resource Development Quarterly* 12 (4): 421–437. doi:10.1002/hrdq.1006.
- Evans, C., and M. Kozhevnikova. 2011. "Styles of Practice: How Learning Is Affected by Students' and Teachers' Perceptions and Beliefs, Conceptions and Approaches to

- Learning." *Research Papers in Education* 26 (2): 133–148. doi:[10.1080/02671522.2011.561973](https://doi.org/10.1080/02671522.2011.561973)
- Flaiss, J. L. 1971. "Measuring Nominal Scale Agreement among Many Raters." *Psychological Bulletin* 76 (5): 378–382. doi:[10.1037/h0031619](https://doi.org/10.1037/h0031619).
- Garet, M., A. Porter, L. Desimone, B. Birman, and K. Yoon. 2001. "What Makes Professional Development Effective? Results from a National Sample of Teachers." *American Educational Research Journal* 38 (3): 915–945. doi:[10.3102/00028312038004915](https://doi.org/10.3102/00028312038004915).
- Geijsel, F., P. Sleegers, R. Van den Berg, and G. Kelchtermans. 2001. "Conditions Fostering the Implementation of Large-scale Innovation Programs in Schools: Teachers' Perspectives." *Educational Administration Quarterly* 37 (1): 130–166. doi:[10.1177/00131610121969262](https://doi.org/10.1177/00131610121969262).
- Gosling, D. 2000. "Using Habermas to Evaluate Two Approaches to Negotiated Assessment." *Assessment and Evaluation in Higher Education* 25 (3): 293–304. doi:[10.1080/02602930050135158](https://doi.org/10.1080/02602930050135158).
- Guay, F., R. J. Vallerand, and C. Blanchard. 2000. "On the Assessment of Situational Intrinsic and Extrinsic Motivation: The Situational Motivation Scale (SIMS)." *Motivation and Emotion* 24 (3): 175–213. doi:[10.1023/A:1005614228250](https://doi.org/10.1023/A:1005614228250).
- Hattie, J., and H. Timperley. 2007. "The Power of Feedback." *Review of Educational Research* 77 (1): 81–112. doi:[10.3102/003465430298487](https://doi.org/10.3102/003465430298487).
- Heller, J. I., K. Sheingold, and C. M. Myford. 1998. "Reasoning about Evidence in Portfolios: Cognitive Foundations for Valid and Reliable Assessment." *Educational Assessment* 5 (1): 5–40. doi:[10.1207/s15326977ea0501_1](https://doi.org/10.1207/s15326977ea0501_1).
- Heller, J. I., K. R. Daehler, N. Wong, M. S. Shinohara, and L. W. Miratrix. 2012. "Differential Effects of Three Professional Development Models on Teacher Knowledge and Student Achievement in Elementary Science." *Journal of Research in Science Teaching* 49 (3): 333–362. doi:[10.1002/tea.21004](https://doi.org/10.1002/tea.21004).
- In de Wal, J., P. J. den Brok, J. G. Hooijer, and R. L. Martens. 2014. "Teachers' Engagement in Professional Learning: Exploring Motivational Profiles." *Learning and Individual Differences* 36: 27–36. doi:[10.1016/j.lindif.2014.08.001](https://doi.org/10.1016/j.lindif.2014.08.001).
- Janssen, F., H. Westbroek, and W. Doyle. 2015. "Practicality Studies: How to Move from What Works in Principle to What Works in Practice." *Journal of the Learning Sciences* 24 (1): 176–186. doi:[10.1080/10508406.2014.954751](https://doi.org/10.1080/10508406.2014.954751).
- Kreber, C., and P. A. Cranton. 1997. "Teaching as Scholarship: A Model for Instructional Development." *Issues and Inquiry in College Learning and Teaching* 19 (2): 4–13.
- Kreber, C., and P. A. Cranton. 2000. "Exploring the Scholarship of Teaching." *Journal of Higher Education* 71: 476–495. doi:[10.2307/2649149](https://doi.org/10.2307/2649149).
- Kyndt, E., F. Dochy, and H. Nijs. 2009. "Learning Conditions for Non-formal and Informal Workplace Learning." *Journal for Workplace Learning* 21 (5): 369–383. doi:[10.1108/13665620910966785](https://doi.org/10.1108/13665620910966785).
- Lillejord, S., and K. Børte. 2019. "Trapped between Accountability and Professional Learning? School Leaders and Teacher Evaluation." *Professional Development in Education* 1–18. doi:[10.1080/19415257.2019.1585384](https://doi.org/10.1080/19415257.2019.1585384).
- Louws, M. L., J. A. Meirink, K. van Veen, and J. H. van Driel. 2017. "Exploring the Relation between Teachers' Perceptions of Workplace Conditions and Their Professional Learning Goals." *Professional Development in Education* 43 (5): 770–788. doi:[10.1080/19415257.2016.1251486](https://doi.org/10.1080/19415257.2016.1251486).
- Lustick, D. 2011. *Certifiable: Teaching, Learning and National Board Certification*. New York: Rowman and Littlefield Education.
- Lustick, D., and G. Sykes. 2006. "National Board Certification as Professional Development: What are Teachers Learning?" *Education Policy Analysis Archives*. 14: 5. Retrieved 10 August 2006 from <http://epaa.asu.edu/epaa/v14n5>

- Lynch, R., P. Mannix McNamara, and N. Seery. 2012. "Promoting Deep Learning in a Teacher Education Program through Self- and Peer-assessment and Feedback." *European Journal of Teacher Education* 35 (2): 179–197. doi:10.1080/02619768.2011.643396.
- Mann, K., J. Gordon, and A. MacLeod. 2009. "Reflection and Reflective Practice in Health Professions Education: A Systematic Review." *Advances in Health Sciences Education, Theory and Practice* 14 (4): 595–621. Doi:10.1007/s10459-007-9090-2.
- May, T. 2013. "Identifying the Characteristics of Written Formative Feedback Used by Assessors in Work-based Qualifications." *Journal of Vocational Education and Training* 65 (1): 18–32. doi:10.1080/13636820.2012.727855.
- McHugh, M. L. 2012. "Interrater Reliability: The Kappa Statistic." *Biochemia Medica* 22 (3): 276–282. doi:10.11613/issn.1846-7482.
- Meirink, J. A., P. C. Meijer, and N. Verloop. 2007. "A Closer Look at Teachers' Individual Learning in Collaborative Settings." *Teachers and Teaching: Theory and Practice* 13 (2): 145–164. doi:10.1080/13540600601152496.
- Oosterbaan, A. E., M. F. Van der Schaaf, L. K. J. Baartman, and K. Stokking. 2010. "Reflection during Portfolio-based Conversations." *International Journal of Educational Research* 49 (4): 151–160. doi:10.1016/j.ijer.2011.02.001.
- Penuel, W. R., B. J. Fishman, R. Y. Yamaguchi, and L. P. Gallagher. 2007. "What Makes Professional Development Effective? Strategies that Foster Curriculum Implementation." *American Education Research Journal* 44 (4): 921–958. doi:10.3102/0002831207308221.
- Popham, W. J. 2006. "All about Accountability/phony Formative Assessments: Buyer Beware." *Educational Leadership* 64 (3): 86–87.
- Porter, A. C., P. Youngs, and A. Odden. 2001. "Advances in Teacher Assessment and Their Uses." In *Handbook of Research on Teaching*, edited by V. Richardson, 259–297. 4th ed. Washington, D.C.: AERA.
- Pratt, M. W., and K. M. Savoy-Levine. 1998. "Contingent Tutoring of Long-division Skills in Fourth and Fifth Graders: Experimental Tests of Some Hypotheses about Scaffolding." *Journal of Applied Developmental Psychology* 19 (2): 287–304. doi:10.1016/S0193-3973(99)80041-0.
- Ross, J. A., and C. D. Bruce. 2007. "Teacher Self-assessment: A Mechanism for Facilitating Professional Growth." *Teaching and Teacher Education* 23 (2): 146–159. doi:10.1016/j.tate.2006.04.035.
- Ruiz-Primo, M. A., and E. M. Furtak. 2007. "Exploring Teachers' Informal Formative Assessment Practices and Students' Understanding in the Context of Scientific Inquiry." *Journal of Research in Science Teaching* 44 (1): 57–84. doi:10.1002/(ISSN)1098-2736.
- Runhaar, P., and K. Sanders. 2013. "Implementing Human Resources Management (HRM) within Dutch VET Institutions: Examining the Fostering and Hindering Factors." *Journal of Vocational Education & Training* 65 (2): 236–255. doi:10.1080/13636820.2013.783612.
- Ryan, R. M., and E. L. Deci. 2000. "Self-determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-being." *American Psychologist* 55: 68–78. doi:10.1037/0003-066X.55.1.68.
- Sadler, D. R. 1989. "Formative Assessment and the Design of Instructional Systems." *Instructional Science* 18 (2): 119–144. doi:10.1007/BF00117714.
- Sato, M., R. C. Wei, and L. Darling-Hammond. 2008. "Improving Teachers' Assessment Practices through Professional Development: The Case of National Board Certification." *American Educational Research Journal* 45 (3): 669–700. doi:10.3102/0002831208316955.
- Schoenfeld, A. H. 1998. "Toward a Theory of Teaching-in-context." *Issues in Education* 4 (1): 1–94. doi:10.1016/S1080-9724(99)80076-7.
- Schön, D. 1983. *The Reflective Practitioner. How Professionals Think in Action*. London: Temple Smith.
- Shulman, L. S. 1986. "Those Who Understand: Knowledge Growth in Teaching." *Educational Researcher* 15: 4–14. doi:10.3102/0013189X015002004.

- Shulman, L. S. 1987. "Knowledge and Teaching: Foundations of the New Reform." *Harvard Educational Review* 57 (1): 1–22. doi:10.17763/haer.57.1.j463w79r56455411.
- Stiggins, R., and J. Chappuis. 2005. "Using Student-involved Classroom Assessment to Close Achievement Gaps." *Theory into Practice* 44 (1): 11–18. doi:10.1207/s15430421tip4401_3.
- Thompson, D. N., and H. K. Burns. 2008. "Reflections: An Essential Element of Evidence-based Practice." *Journal of Emergency Nursing* 34 (3): 246–248. doi:10.1016/j.jen.2008.02.024.
- Tigelaar, D., P. Sins, and J. van Driel. 2017. "Fostering Students' Reflection: Examining Relations between Elements of Teachers' Knowledge." *Research Papers in Education* 32 (3): 353–375. doi:10.1080/02671522.2016.1225790.
- Uhlenbeck, A. M., N. Verloop, and D. Beijaard. 2002. "Requirements for an Assessment Procedure for Beginning Teachers: Implications from Recent Theories on Teaching and Assessment." *Teacher College Record* 104 (2): 242–272. doi:10.1111/1467-9620.00162.
- Vallerand, R. J. 1997. "Toward a Hierarchical Model of Intrinsic and Extrinsic Motivation." In *Advances in Experimental Social Psychology*, edited by M. P. Zanna, 271–360. Vol. 29. San Diego: Academic.
- Van de Pol, J., M. Volman, and J. Beishuizen. 2011. "Patterns of Contingent Teaching in Teacher-student Interaction." *Learning and Instruction* 21 (1): 46–57. doi:10.1016/j.learninstruc.2009.10.004.
- Van den Bergh, L., A. Ros, and D. Beijaard. 2014. "Improving Teacher Feedback during Active Learning: Effects of a Professional Development Program." *American Educational Research Journal* 51 (4): 772–809. doi:10.3102/0002831214531322.
- Van Diggelen, M., P. J. den Brok, and D. Beijaard. 2013. "Teachers' Use of a Self-assessment Procedure: The Role of Criteria, Standards, Feedback and Reflection." *Teachers and Teaching* 19 (2): 115–134. doi:10.1080/13540602.2013.741834.
- Van Driel, J. H., and A. Berry. 2010. "The Teacher Education Knowledge Base: Pedagogical Content Knowledge." In *International Encyclopedia of Education*, 3rd. Vol. 7. B. McGraw, P. L. Peterson, and E. Baker, edited by. 656–661. Oxford:Elsevier.
- Verberg, C. P. M., D. E. H. Tigelaar, and N. Verloop. 2013. "Teacher Learning through Participation in a Negotiated Assessment Procedure." *Teachers and Teaching: Theory and Practice* 19 (2): 172–187. doi:10.1080/13540602.2013.741842.
- Verloop, N., J. H. Van Driel, and P. C. Meijer. 2001. "Teacher Knowledge and the Knowledge Base of Teaching." *International Journal of Educational Research* 35 (5): 441–461. doi:10.1016/S0883-0355(02)00003-4.
- Vygotsky, L. S. 1978. *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, MA: Harvard University press.