Assessment for Learning: developing the required teacher competencies

Christel Wolterinck, Cindy Poortman, Kim Schildkamp & Adrie Visscher

Published online: 19 Sep 2022.

Submit your article to this journal ⍩

Article views: 698

View related articles ⍩

View Crossmark data ⍩

Routledge
Taylor & Francis Group

ARTICLE

Assessment for Learning: developing the required teacher competencies

Christel Wolterinck, Cindy Poortman ⓘ, Kim Schildkamp and Adrie Visscher ⓘ

Elan, Department of Teacher Development Faculty of Behavioural, Management, and Social Sciences of the University of Twente, Enschede, The Netherlands

ABSTRACT

Assessment for Learning (AfL) is complex requiring teacher compe- tences for collecting, analysing and interpreting evidence of stu- dent progress, and for adapting instruction and learning accordingly. In this study, the four-component instructional design (4C/ID) model for complex learning was used to design a teacher professional development (TPD) programme for developing tea- chers' AfL competences. The evaluation of the implementation of the program showed that teachers were positive about the rele- vance and usability of the program, and about the AfL knowledge, skills and attitudes gained. This study yields important lessons about how to support teachers in learning to apply AfL in their classrooms. The study shows that the 4C/ID model can provide the learning-psychological basis for the design of TPD programmes for the acquisition of complex teacher skills, that require the integra- tion of skills, knowledge and attitudes.

Introduction

Assessment for Learning (AfL) is a conscious attempt to make (classroom) assessment an essential part of teaching and learning (Wiliam 2011). The information derived from different assessment sources can be used as a form of continuous feedback to steer learning (Hattie and Timperley 2007; Wiliam and Leahy 2015). Teachers are increasingly expected to have the skills, knowledge and attitudes needed to enact AfL in their class- room (Heitink et al. 2016). Although AfL can enhance student learning (Black and Wiliam 1998), research has shown that the implementation of AfL often is not effective (Bennett 2011). A possible explanation of the lack of positive effects could be that teachers may lack the professional competences needed to implement AfL in their classroom effectively (Schildkamp et al. 2020).

Competences related to collecting, analysing and interpreting evidence from assess- ment, and adapting instruction and learning accordingly are crucial for the effective implementation of AfL in classroom practice (Lee 2011). Research on teachers' use of AfL has indicated room for improvement when it comes to these competences (Baartman and Gulikers 2017; Kippers et al. 2018). Therefore, investment in professional development

CONTACT Christel Wolterinck ✉ c.h.d.wolterinck@utwente.nl ⬚ Faculty of Behavioural, Management, and Social Sciences of the University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands

(PD) is crucial (Heitink et al. 2016). To effectively support teachers in using AfL, we developed and implemented the AfL teacher professional development (AfL-TPD) program using the Four-Component Instructional Design (4C/ID) model (Van Merriënboer and Kirschner 2017). In the design of teacher professional development programs an explicit, validated, learning-psychological basis, that takes into account what it requires to acquire knowledge, skills and attitudes, is not seldom missing. The 4C/ID model can provide such a basis.

The 4C/ID model is a task-centred approach suitable for teaching complex professional tasks that require the integration of knowledge, skills, and attitudes, and the coordination of various skills (Van Merriënboer and Dolmans 2007). The aim of this study was to evaluate the effects of the AfL-TPD program that was developed based on the following research questions:

RQ1: How relevant and usable did teachers find the assessment for learning professional development program that was designed on the basis of the four components instructional design model?

RQ2: What knowledge, skills and attitudes needed for AfL did teachers learn as a result of participating in the AfL professional development program?

Theoretical framework

Assessment for learning

AfL is an approach to formative assessment that takes place as part of ongoing classroom practice and is meant to enhance the quality of students' learning processes (Black and Wiliam 1998; Klenowski 2009; Stobart 2008). Wiliam (2011) distinguished five core strategies for implementing AfL in classroom practice: 1. clarifying, sharing and understanding learning goals and criteria for success; 2. engineering effective classroom discussions, activities, and learning tasks that elicit evidence of learning; 3. providing feedback that moves learning forward; 4. encouraging learners to be instructional resources for one another; 5. encouraging learners to be owners of their own learning. Studies focusing on the implementation of the five strategies of AfL in teaching practices emphasise the cyclical character of the AfL process and emphasise that the collected data are analysed (turning data into information) and transformed into decisions, so that teachers can provide feedback for students (Ruiz-Primo and Furtak 2007; Baartman and Gulikers 2017; Schildkamp et al. 2020). AfL is a complex teacher competence, and TPD should explicitly support teachers in integrating AfL strategies coherently in their classroom practice, in order to maximise its potential impact (Lee 2011).

Design principles for the AfL-professional development program

From cognitive psychology, we know that the following is important for learning complex tasks and to ensure the transfer of learning: whole meaningful, authentic and varied learning tasks; ordering learning tasks from simple to complex, in combination with the gradual decrease in learner support and distinguishing between the non-routine and routine aspects of complex skills (Van Merriënboer, Kester, and Paas 2006). The AfL-TPD program in this study was designed using a task-centred instructional design model, the

Four-Component Instructional Design (4C/ID) model, in which the design of authentic learning tasks follows an holistic approach dealing with real-life professional tasks that allow the learner to practice all the nonroutine and routine aspects of a complex professional task (AfL in this study) simultaneously (Van Merriënboer 2007). The 4C/ID model includes four interrelated components: a sequence of *learning tasks* based on authentic professional tasks, *supportive information* describing how to approach the tasks and how the domain is organised, *procedural information* describing step-by-step procedures to perform routine aspects of the tasks, and *part-task practice* for repetition of aspects that need to be highly automated (Van Merriënboer and Dolmans 2007; Van Merriënboer, Kirschner, and Kester 2017). The learning tasks for teachers which need to be conducted in the AfL-TPD program are based on the results of a Cognitive Task Analysis (CTA; Clark et al. 2008) of teachers who are AfL-experts.

The CTA was conducted in a previous study which provided an overview of all constituent skills and the relationships between these skills, visualised in a skills hierarchy depicted in Figure 1 (based on Wolterinck et al. 2019). Four main skills make up AfL and are closely interrelated: the teacher prepares a lesson (1), based on the evaluation of the previous lesson (2) and based on the preparation of the set of lessons for an instructional topic (3). This enables the teacher to enact the lessons while tailoring instruction to students' learning needs, eliciting evidence of learning and encouraging students to be owners of their learning (4). Learning and performing these four main skills is facilitated by a subset of constituent skills (i.e. those positioned next to the main four skills in Figure 1). These constituent skills are depicted vertically, implying that they must be performed in order.

Besides the skills hierarchy, the CTA revealed the kinds of teacher knowledge are critical for using AfL successfully in classrooms (Heritage 2007; Wolterinck et al. 2019): domain knowledge, pedagogical content knowledge, knowledge of students' previous learning, and assessment knowledge. Teachers need to know how to align formative assessment with instructional goals, which requires knowledge of students' previous learning in relation to the curriculum and the learning goals and knowledge of students' misconceptions. Teachers must also know a range of formative assessment strategies, and know how to apply them, to maximise the opportunities for gathering evidence, while understanding that the quality of the assessment is an important concern. Finally, teachers need to know that AfL requires a joint effort by teachers and students: students' self and peer assessments provide important opportunities for establishing their current learning status. With regard to teacher attitude: teachers need to view AfL as a worthwhile process that yields valuable and actionable information about students' learning, and they need to view AfL and the teaching process as inseparable, recognising that one cannot happen without the other (Heritage 2007; Wolterinck et al. 2019).

The CTA included an analysis of the factors that influence the complexity of implementing AfL: student group composition (e.g. number of students, variation in student levels), the content of the lesson (goal and topic), curriculum material (suggestions for remediation and materials or not) and support from the school (collaboration and facilities) (Wolterinck et al. 2019). These factors were used as focal points for selecting learning tasks for the AfL-TPD program according to their complexity.

The 4C/ID design guidelines led to several design decisions. First, the main part of the program took place within schools, using the w*hole-task approach* with learning tasks that
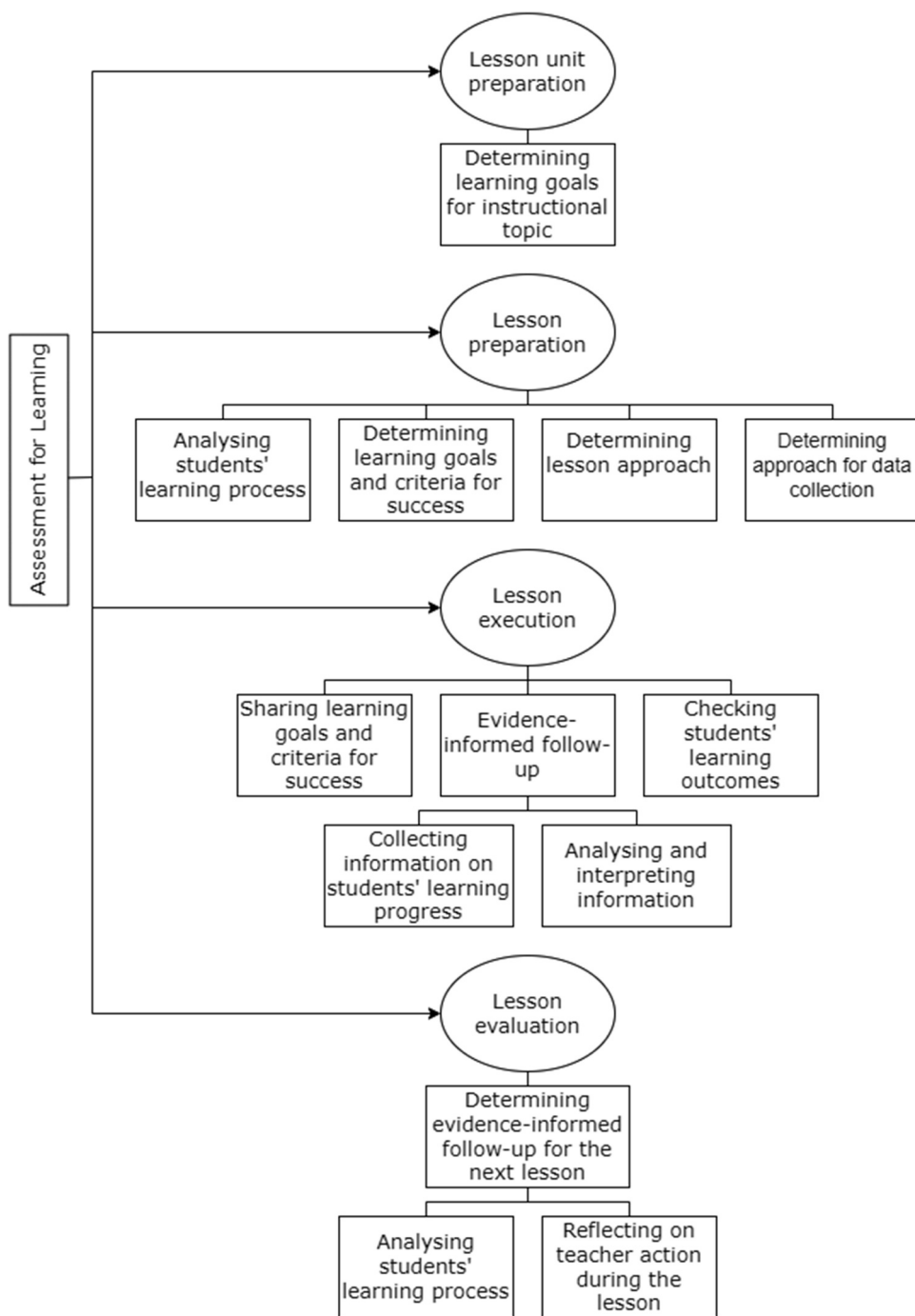
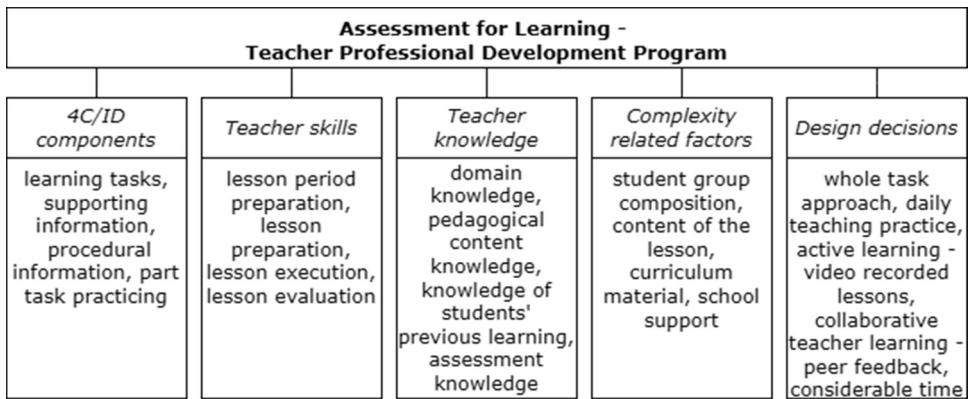**Figure 1.** Skills required for assessment for learning.

Figure 2. Essential building blocks of AfL-TPD program.

represent the range of possible real-life tasks the teacher may encounter. This means that teachers practiced AfL within the full instructional teaching sequence of preparing a lesson series for an instructional topic, preparing lessons, enacting those lessons and evaluating the lessons, in which all constituent skills are supposed to be present. Second, the program provided many opportunities for practice and experimentation, by applying newly acquired skills during *daily teaching practice*. The program focused on relevant content, highlighting relevant domain knowledge, and teacher learning is activated by connecting what teachers already know or can do with what has to be learned. Third, the program stimulated *active learning using video-recorded lessons*, demonstrating expert strategies and stimulating reflective skills. Fourth, the program stimulated *collaborative teacher learning*, activating teachers as peers in the feedback process. Finally, the program spans approximately 8 months, to provide *ample time* to reach the desired outcomes. Figure 2 shows a schematic overview of the essential building blocks of the AfL-TPD program.

Evaluation of AfL-professional development program

Merrill's first principles of instruction (Merrill 2012) provide a system for assessing instruction for complex learning of authentic, real-world, whole tasks. These principles are based on a synthesis of instructional design theories, consistent with the 4C/ID model of instructional design, and therefore are suitable for a formative evaluation of the *relevance and usability* of the AfL-TPD program (Frick et al. 2010; Merrill 2012). The five criteria that 4C/ID-designed instruction must meet are: (1) instruction is problem- or task-centred, (2) teacher learning is activated by connecting what teachers already know or can do, with what is to be learned, (3) teachers are exposed to demonstrations of what they are to learn, (4) teachers have opportunities to try out what they have learned, in combination with instructor coaching and feedback, and (5) teachers integrate what they have learned into their own classroom practices (Merrill 2012).

For assessing teachers' *learning* in the AfL-TPD program, teacher skill improvement was examined in terms of the constituent skills described in the skills hierarchy and the related

performance objectives. For each of the constituent skills, standards were set for desired performance after completing the program (Appendix A). To examine changes in teacher *knowledge* and *attitudes*, the following elements of teacher knowledge and attitude, based on the CTA, it was evaluated whether teachers: knew how to align formative assessment with instructional goals; had knowledge about a range of formative assessment strategies; had knowledge of students' previous learning in relation to the curriculum and learning goals; had knowledge of students' misconceptions; viewed AfL as a worthwhile process that yields valuable and actionable information about students' learning; viewed AfL and the teaching process as inseparable and recognised that one cannot happen without the other.

Method

This study was conducted in the context of Dutch secondary education (students 12–18 years old). To study teachers' satisfaction with the AfL-TPD program in regard to the 4C/ID principles and teachers' learning in terms of their knowledge, skills, and attitude necessary for implementing AfL in practice, we adopted a mixed-methods methodology. Teachers completed a questionnaire and also participated in focus group interviews for triangulation purposes. Coaches from the AfL-TPD program participated in a separate focus group interview (Cohen, Manion, and Morrison 2013).

Assessment for learning – professional development program

The AfL-TPD program was conducted on a subject-specific basis (English language, Dutch language and chemistry) and had a study load of approximately 40 hours spread over 8 months. This included 20 hours of contact time, spent in 4 meetings, and 20 hours of practicing time, consisting of carrying out learning tasks. In the AfL-TPD program we focused on workplace-based learning and used participants' lesson series and their daily lessons (for the participants their English language, Dutch language and chemistry lessons) as learning tasks. These real-life whole tasks present the full range of variability, require all the necessary skills, and allow daily opportunities for practice. Figure 3 shows a schematic overview of the AfL-TPD program and appendix B presents the final blueprint for the design of the introductory meeting at the start of the program.

In the 4C/ID model, learning tasks at the beginning of the program are ideally practiced in a safe learning environment in which errors do not have serious consequences and where support and feedback can be easily presented just in time (Van Merriënboer and Kirschner 2017). Because daily lessons were used as learning tasks in the AfL-TPD program, it was not possible to manipulate the complexity-related factors identified in the CTA (e.g. student diversity, group size). In the current design it was for this reason not possible to work with task classes in which each task class includes a set of equivalent learning tasks that are at the same level of complexity. This problem was counteracted by using an *emphasis manipulation* approach, to sequence learning tasks. In this approach, the learner's attention is actively allocated to an aspect of a certain skill or task (Gopher 2006), for example, task class 1 focused on sharing learning goals and criteria for success during the lesson (see appendix C for a more extensive description of task class 1).
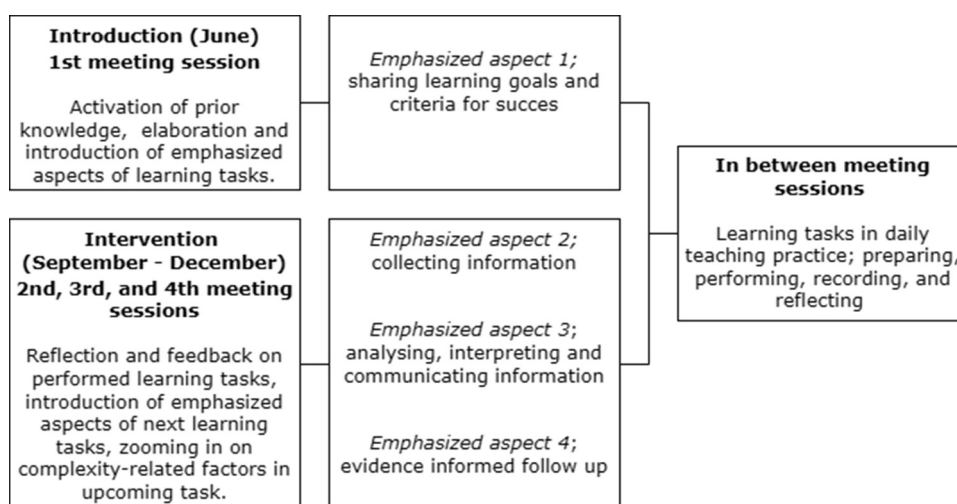
Figure 3. Overview AfL-TPD program.

To compensate for the reduced support during the learning tasks in daily lessons, we increased support before and after the learning tasks (Frerejean et al. 2021). The first face-to-face meeting was planned to prepare teachers for independent practice in their own classrooms focusing on sharing learning goals and criteria for success. The coach guided the teachers in studying exemplars and clarifying expectations regarding what good practices look like (Carless and Chan 2017). For six weeks, teachers practiced AfL within the full instructional teaching sequence and had to prepare a lesson series for an instructional topic, prepare lessons, enact those lessons and evaluate the lessons. Teachers learned how to apply the whole cycle of AfL, while actively focusing their attention on sharing learning goals and criteria for success during reflection and feedback activities. By the end of the program, all aspects of the task had been addressed. The final task class ended with designing a lesson series; in the fourth and final meeting session, school teams shared their experiences and reflected on their overall improvement. This concluded the 8-month AfL-TPD program.

Participants

Teachers

Four schools voluntarily signed up to participate in the AfL-TPD program, responding to a call communicated via a newsletter sent out by one of the largest school boards in the Netherlands. These schools were each asked to have three teachers participate, preferably one English language teacher, one Dutch language teacher and one chemistry teacher. We started the program with 12 teachers in total, of which, two participants left the program after the second meeting, because the AfL-TPD program was too demanding and too hard for them to combine with their own teaching job. Eight out of ten teachers in the program responded to the questionnaire and participated in the focus groups. For an overview of the key characteristics of the ten teachers see Table 1.

Table 1. Key characteristics of respondents.

		N
Subject	English language	2
	Dutch language	4
	Chemistry	4
Gender	Female	9
	Male	1
Teaching experience	< 5 years	1
	5–14 years	5
	15–24 years	2
	> 25 years	2

Coaches

A design team consisting of the researchers of this paper, and coaches from the Dutch national expertise centre for curriculum design set out to design and implement the AfL-TPD program. The coaches for the AfL-TPD program were subject-matter specialists, one English language, one Dutch language, and one chemistry specialist, and they were experienced in supporting teachers in secondary education. One of the coaches, the English language specialist, had ten years of experience with supporting teachers in using AfL. The other two coaches were less familiar with supporting teachers in using AfL, and were therefore trained by the English language specialist for the AfL-TPD program. The program meetings were carefully prepared by the design team.

Data collection

Questionnaire

In this explorative study a questionnaire was used for evaluating the AfL-TPD program in which participants' satisfaction and whether teachers found the program relevant and usable (utility of the content, timeliness, materials and support, organisation as well as satisfaction with the program design) was examined using 27 statements. The statements also covered various other aspects of the program, including the materials and support provided and the organisation of the meetings. To optimise its face validity, the instrument was piloted with two expert researchers and four teachers (Cohen, Manion, and Morrison 2013). Based on their comments, adjustments were made to the questionnaire. The questionnaire was administered after the final program meeting. Teachers were asked to express their agreement with the statements (e.g. 'In this program, I was satisfied with the course material that was offered'), using a 5-point Likert scale ranging from 'strongly disagree' to 'strongly agree'. The questionnaire items were in Dutch. In addition, teachers could give a brief explanation of the chosen answer.

focus group sessions with teachers and coaches

Four weeks after the final program meeting, two focus groups with four teachers each were conducted and one focus group session with the three coaches was held to evaluate from multiple viewpoints (Cohen, Manion, and Morrison 2013). Two teachers from each of the four participating schools were invited to attend the focus group sessions for teachers, which resulted in the representation of all three subjects. The goal of these meetings was

to evaluate teachers' satisfaction with the program and teachers' learning outcomes (skills, knowledge and attitude) for triangulation purposes. The first goal of the focus group session for coaches was to have the coaches evaluate teachers' satisfaction regarding various aspects of the AfL-TPD program. The second goal was to have the coaches evaluate teachers' learning outcomes (skills, knowledge and attitude regarding AfL).

All three focus group sessions lasted two hours and were audio-recorded. The teachers and coaches were first asked to respond individually to a series of statements and questions based on the theoretical framework and the questionnaire data, followed by a group discussion yielding a collective rather than an individual view (Cohen, Manion, and Morrison 2013). The statements (e.g. 'I was satisfied with the approach in which new knowledge was integrated into my own daily teaching practice', 'In order to improve the AfL-TPD program, teachers need more examples from practice') and questions (e.g. 'What did you learn in terms of knowledge of AfL?', 'How satisfied do you think teachers were with the guidance and materials?') used in the focus group session were discussed in advance with a researcher, after which the formulation of the statements and the questions was adjusted.

Analysis

The first research question (teacher satisfaction) was partly answered by analysing the frequency distribution of the questionnaire responses. Additionally, both the first and the second research question (teacher learning) were answered by analysing the recordings of the three focus group sessions. Based on the theoretical framework, an a priori coding scheme with 20 codes was developed. All focus group sessions were audio-taped and transcribed, and the inter-rater reliability between the three coders was calculated across ten percent of the focus group data, to increase reliability with other researchers (Poortman and Schildkamp 2012). They found an acceptable Cohen's Kappa of 0.72 (Eggen and Sanders 1993). Table 2 presents the coding scheme. After coding the focus group data, we summarised what teachers or coaches said during the focus group interviews related to each code. Quotes from the respondents were translated into English for use in this article. Based on the analyses of multiple focus group interviews we could report detailed answers to the two research questions (Cohen, Manion, and Morrison 2013).

Table 2. Coding scheme.

Concept	Codes
Satisfaction	Design
	Materials and support
	Organisation
Learning	Knowledge: general, preparing a lesson period; preparing a lesson; lesson execution; lesson evaluation.
	Skill: general, preparing a lesson period; preparing a lesson; lesson enactment; lesson evaluation.
	Attitude: general, preparing a lesson period; preparing a lesson; lesson execution; lesson evaluation.

Results

Teacher satisfaction

Quality of the design

The questionnaire showed that participants were satisfied to very satisfied with the content of the AfL-TPD program. For example, all eight participants indicated contentment with the practice-oriented tasks, the structure offered to link new knowledge with prior knowledge, and the examples used for modelling. In the focus groups, teachers reported their satisfaction with the applicability of what was learned in their own teaching practice.

All eight teachers were positive about being actively engaged by the program during the meetings, for example, via discussing examples and giving and receiving feedback. They also appreciated the assignments between meetings, which were linked to their daily practice. In the focus groups, teachers mentioned their appreciation for the exchange of experiences during the meeting sessions and would like to see this happen more frequently between the meetings, when they are applying newly learned insights in their own practice: 'For example, sending each other videos between training sessions enables you to give each other more feedback. I could have got more out of the training'. Seven out of eight teachers expressed their satisfaction about the structure of the skills hierarchy. It gave them an overview and clarified the cyclical process of AfL. This also applied to the exemplars used, which stimulated discussion and developing a sense of quality together.

Teachers indicated some room for improvement regarding the program, for example, more explicit use of prior knowledge, and providing success criteria for the assignments. Teachers also mentioned differentiation in the AfL-TPD program as an area for improvement, such that teachers receive instruction and tasks tailored to their needs: 'One teacher develops quickly and the other very slowly and there can be more focus on this in the program'. Four teachers were only moderately satisfied with the focus on sharing knowledge with colleagues in their school and how literature was used.

Coaches indicated that the program should focus more on the cyclical process of AfL, because teachers need to understand that AfL is about students' learning processes instead of focusing on students' learning outcomes. Otherwise, teachers tend to use AfL as an instrument for testing without grades: 'It is a process, formative assessment is not just a test to see where the student currently is. It's about them getting further and what is needed to do that'; 'Where do I want to take the students and what can I do as a teacher to make the students aware of their own learning process, that is what I have developed'. Teachers generally came across as satisfied, according to the coaches, as demonstrated by, for example, high turnout at meetings (only one teacher missed a meeting session due to sickness), active participation in discussions, asking questions, and actively working on the assignments in their own teaching practice. One of the teachers indicated to a coach, for example, that after 35 years of teaching, this program had given her job a new boost. With great enthusiasm, she now reported implementing AfL in her lessons.

Materials and support

The questionnaire showed that all eight teachers were very satisfied with the applicability of materials to their own teaching practice, such as the use of exemplars (video fragments of lessons), discussing examples from participants' teaching practices and sharing experiences and teaching materials. In the focus groups, teachers reported that watching each other's videos of lessons was instructive and gave insight into their own learning processes. Using peer feedback to look at one's own actions is easier than directly reflecting on one's own actions'. The book with supportive information was also very practical as a reference work. All teachers experienced a lot of guidance and support from the coaches, because each subject had a curriculum specialist present and the group of teachers was relatively small. Coaches stated that they could almost give one-on-one guidance and many questions could be answered. Teachers also addressed their questions to the coaches between the meetings via email. All eight teachers totally agreed with statements such as: 'Questions were adequately answered by the coaches" and "I am satisfied with the course material that was offered'.

Teachers indicated some areas for improvement, for example, in the second meeting session the teachers received a syllabus containing an overview of information collection techniques as supportive information for task class 2. Instead of just providing the syllabus, it would be better to share the syllabus as linked with an assignment, for example, asking teachers to choose a technique and develop an application of it in their own lesson. Other areas for improvement reported were, for example, practicing more with formulating and sharing learning goals and success criteria. The assignments should be provided with clear formulated success criteria for the participants: 'As participants, we should also set our goals and our success criteria more in these meetings'. Teachers suggested that the coaches could demonstrate greater variety in the exemplars (not just language examples), time for reflection to translate the example to one's own teaching practice, and more focus on the content of the book. The meetings could also be less traditional, meaning that coaches could demonstrate the use of AfL more explicitly in their own teaching, or, as one of the teachers stated, 'teach as you preach'.

Organisation and learning management

The questionnaire showed that teachers were very satisfied with the organisation of the meetings; for example, seven out of eight teachers indicated their contentment with the spread of the meetings over time, the efficiency of the meetings and the facilities used. All teachers also mentioned their contentment about working in small groups (up to 4) of varying composition, as this allowed discussions and the exchange of experiences. Participating with colleagues from the same school was experienced positively, because it stimulated learning between the meetings in teachers' own teaching practice in their own schools, by giving peer feedback and discussing video recorded lessons.

Both teachers and coaches indicated some room for improvement; for example, they suggested taking more time between the last two meetings to build in peer feedback: 'With the feedback obtained, the next lesson can be given and recorded and be compared with the former lesson, in order to make development visible. Learning does not just take place during the meetings, but especially between the meetings when working in your own teaching practice'. Both teachers and coaches also suggested setting up a digital environment for sharing the video fragments and subsequently providing each other with

feedback, and for conducting intake interviews prior to the AfL-TPD program to get an idea of teachers' prior knowledge and experience with AfL. Teachers suggested involving school leaders in the program, because school leaders need to support their teachers in broadening AfL at their own school: 'Have the school leader join you in one of the first meetings. They should be able to support their teachers to broaden the developments around formative assessment in their own schools'.

Teacher learning

Knowledge and skills

Teachers and coaches both indicated that teachers gained more knowledge about AfL and about the context in which AfL can be used in one's own teaching practice. For example, the definition of AfL became clear; teachers were able to use the terminology and see possibilities for applying AfL in their lessons, and teachers indicated that the structure of the skills hierarchy was supportive. The importance of sharing learning goals and criteria for success with the students and the importance of collecting information about student learning based on learning goals are examples of what the focus group teachers said they learned in the program. Another key learning outcome according to the focus group teachers is the awareness of building AfL as a cyclical process into their lessons. Teachers said they learned that AfL is a cyclical process in which teachers must use the strategies coherently together with students, in order to apply it effectively in the classroom: 'AfL is not just a tool for gathering information about students' learning processes. It's about how I can help students to develop further and what is required to do so and also to promote students' ownership',

Teachers reported that the AfL-TPD program improved their skills; for example, they prepared the lessons more consciously by thinking about the necessary prior knowledge, checking this with students and building on it further. One teacher indicated that she now differentiates more in her lessons; at the beginning of the lesson, she checks which students have mastered the content and can continue working on it themselves and which students need extended instruction. Teachers also felt more comfortable about the skill of collecting and analysing information on students' learning processes and taking action in their teaching: 'I've become much better at collecting information, for example, the use of Socrative [cloud-based student response system] and asking questions, and acting on it. It makes my lessons more effective than before'. Both teachers and coaches reported that teachers developed their skills for evaluating the lessons more consciously using the video fragments from their lessons and asking reflective questions such as: 'What does this mean for my next lesson and how can I change the daily grind?'. The program gave them tools, for example, the use of exit cards to reflect on one's own teaching practice: 'As you are constantly collecting information about students, you can also see more about the effects of your own actions'.

The AfL-TPD program helped teachers to start applying and experimenting with AfL in their own practice: 'I'm noticing how AfL is changing my lessons and that I approach things differently because I'm more aware of how to achieve more'; '(...) you get to know your students better'. Coaches also indicated that teachers showed their AfL skills in several good examples: 'One participant developed and taught lessons about organic chemistry in which she determined and shared the learning goals together with her

students and then collected information about students' learning based on the learning goals'. One teacher indicated that her knowledge and skills had certainly increased; for example, when preparing the lesson, she is now more conscious of students' prior knowledge, what they should know and be able to do, and how she can act on that.

Although teachers indicated that they gained more knowledge and skills related to AfL, both teachers and coaches indicated that more time is needed to practice it properly, for it to become a daily routine: 'I have to stimulate myself to let it become a fixed pattern; One good example in the program does not mean that teachers are able to apply AfL in every lesson'. According to both teachers and coaches, the strategy of 'formulating learning goals and criteria for success' was still challenging and should be given more attention in the program. At the start, teachers indicated themselves to be skilled at determining learning goals and success criteria, but gradually it turned out to be more complex than teachers thought: 'Teachers focus too much on details (little twigs) and not enough on the higher learning goal, the bigger picture (tribe). For example, the correct conjugation of the English verbs is not a learning objective, but is a success criterion for the correct writing of a business mail'.

Attitude

The AfL-TPD program succeeded in changing teachers' attitude towards the use of AfL in their teaching practice. Teachers said they learned to see AfL as a worthwhile process that yields valuable and actionable information about students' learning, and teachers recognised that AfL and teaching are closely related. They also indicated that the program had taken them out of their routine and made them look more critically at their own teaching practice: 'I have become more motivated and positive about AfL, because I started to work with it in my class and it improves the quality of my lessons'. Participating in the AfL-TPD program raised awareness that things need to change in secondary education: less use of summative assessments and less whole classroom teaching. Coaches confirmed the change in teachers' attitude towards AfL: 'Teachers think more consciously about AfL and are more enthusiastic about the use of AfL together with students'. Coaches reported that teachers were more aware of the importance of setting and sharing the learning goals for the lesson together with the students, the importance of focusing on students' learning processes instead of their learning outcomes, and when to apply AfL strategies in their lessons.

Conclusion and discussion

The study yields important lessons about how to support teachers in learning to use AfL in their teaching practice. More in general, the study shows that the 4C/ID model (Van Merriënboer and Kirschner 2017) can provide the learning-psychological basis for the design of professional development programs for the acquisition of complex teacher skills, that require the integration of skills, knowledge and attitudes.

Contentment with relevance and usability of the AfL-TPD program, and teachers' knowledge, skills, and attitudes related to AfL gained

The evaluation of the implementation of the AfL-TPD program showed that teachers were positive about their active engagement during the meetings, developing a sense of quality together via discussing exemplars and giving and receiving feedback on video recorded lessons. Teachers reported that the content of the AfL-TPD program was highly applicable, using learning tasks in their own teaching practice. The results show that teachers developed their idea of AfL and related concepts and of the contexts in which AfL can be used in their own teaching practice. Teachers reported learning that AfL is a cyclical process, instead of just an instrument for testing without grades, in which teacher must use the strategies coherently together with students, in order to apply it effectively in their own classroom. Teachers became more skilled in analysing students' learning processes and in reflecting on their own actions.

Important design decisions for implementing an AFL-TPD program

The program was developed using the 4C/ID model, including a *whole task approach* with real-live learning tasks, decreasing scaffolding, and distinguishing between non-routine and routine aspects of the complex skill (Van Merriënboer, Kester, and Paas 2006). Using a whole-task approach can help teachers transferring newly learned skills to their professional task, their daily lessons in their own classrooms (Van Merriënboer and Kester 2015), and teachers this way experience the complexity of implementing AfL while they practice. Although the design team had little control over task sequence or complexity, because daily lessons were used as learning tasks, this problem was counteracted by using tasks with a high amount of support and guidance during the face-to-face meetings.

Second, the program provided many opportunities for practice and experimentation by applying newly acquired skills in teachers' *own teaching practice* and teachers were actively engaged by means of practice-oriented tasks in which learning was stimulated, by connecting what they already knew or could do with what they had learned. Teachers were asked to practice continuously between the face-to-face meetings. This distribution of practice opportunities is considered important to improve long-term learning, which can be explained by the spaced-practice principle: 'practice opportunities distributed across time rather than massed within one session' (Petersen-Brown et al. 2019, 977).

Third, the program stimulates active learning using *exemplars and video recorded lessons*, demonstrating expert strategies and stimulating reflective skills. An essential component of the program, encouraging teachers to reflect, was the dialogic use of exemplars (Carless and Chan 2017). The analysis of exemplars was a powerful way of developing teachers' understanding of the nature of quality by producing accounts of strengths, weaknesses and how the task could have been done better.

Fourth, the use of modelling examples and video fragments from their own teaching practice facilitated the discussions between peers (in small groups) about the performance objectives of AfL and facilitated the process of giving and receiving *feedback* (Lynch, McNamara, and Seery 2012). The use of these reflective discussions created a rich learning environment for teachers and facilitated *collaborative teacher learning* (Vangrieken et al. 2015). Also, a *team* of three teachers working in the same school

attending the AfL-TPD program, which stimulated opportunities for feedback and reflection between the meetings on their own teaching practice by using these colleagues as peers (Smith 2016). Finally, the program spanned approximately *8 months*, which could provide ample time to develop the basic teacher competences needed for AfL. More time is needed to further develop these competences in their own teaching practice (e.g. Darling-hammond, Hyler, and Gardner 2017).

Improvement suggestions for the AfL-TPD program

The experiences of teachers and coaches with the AfL-TPD program were predominantly positive, although suggestions for improvement were also mentioned (Figure 4). First, the use of a video-based professional learning platform should support providing peer-to-peer feedback to allow reflective discussions online between the meeting sessions, and to engage teachers as learning resources for one another. After all, there is strong scientific evidence that feedback can enhance learning (Hattie and Timperley 2007; Kluger and DeNisi 1996). Second, coaches could provide more variety in the modelling examples used (e.g. not just videos from language lessons), and stimulate effective translation of the examples to other subject domains. Therefore, teachers need curriculum content knowledge and pedagogical content knowledge to make the transfer to their own subject matter (Mandinach and Gummer 2016). Third, the coaches could demonstrate the use of AfL more explicitly in their teaching and assignments; for example, the program could be more targeted by determining success criteria for the assignments in advance, and by developing an understanding of the features of quality through discussing exemplars (Carless and Chan 2017; Hamodi, López-Pastor, and López-Pastor 2017). The coaches should not operate on a one-size-fits-all basis, but should differentiate professional development activities deliberately so that teachers receive instruction matching their needs (George 2005). Finally, teachers suggested involving school leaders in the AfL-TPD program, because school leaders need to support their teachers in broadening AfL at their own school (Heitink et al. 2016; Schildkamp et al. 2020).
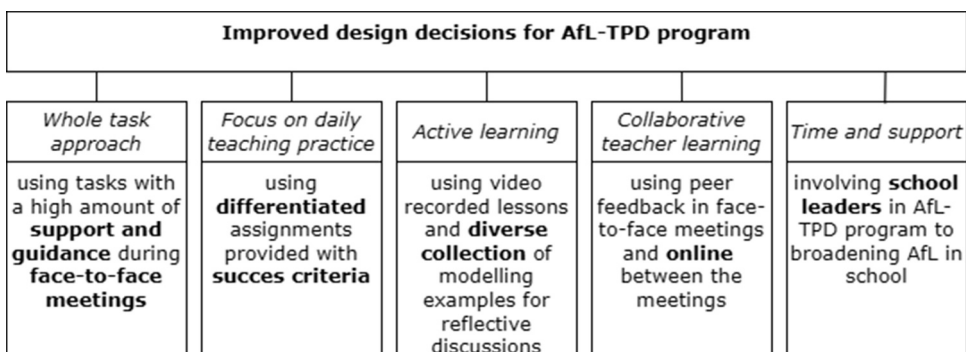

Figure 4. Improved design decisions for AfL-TPD program.

Limitations and implications for practice and research

We acknowledge several limitations of our study. First, it is important to emphasise that in this study, a small group of 10 motivated teachers participated in the AfL-TPD program and developed their competences. The context for future professional development programs might be different; for example, scaling up to larger groups might result in less motivated participants, less involved coaches and result in less successful implementation of AfL in teaching practices. Second, teacher self-report data were collected by means of a questionnaire and focus group interviews. We analysed the effects on teacher learning by using teachers' and coaches' perceptions, which could have been biased; for example, the teachers and coaches in the program might have been inclined to respond positively because they had put a lot of effort into it. Although self-reports have disadvantages, in this qualitative study teachers can provide insight into their own intentions, thought processes, knowledge and beliefs, and therefore it seems important to involve teachers in the evaluation process (Goe, Bell, and Little 2008).

In this small-scale study, we used 4C/ID for the design of an AfL-TPD program. This use of specific design principles for a program provides valuable information and insights for teacher professional development. For example, the dialogic use of exemplars or the use of video-recorded fragments of lessons to stimulate reflective discussions on teaching practice can become a routine within school teams and stimulate a learning culture within the school. The design principles illustrated in this study can serve as an example for other research studies and programs concerning professional development for complex teacher skills.

Despite the fact that teachers became more knowledgeable about AfL, they indicated that applying AfL was not yet a daily routine. Teachers pointed to the pitfall of easily falling back on their routines, due to the hectic pace of the school day. The AfL-TPD program enabled teachers to start applying AfL in their teaching practice, but learning and coaching should continue in the school, supported by students, colleagues and school leaders (Schildkamp et al. 2020).

Further research could examine the effects of this AfL-TPD program on teachers' use of AfL in classroom practice and student achievement, using a larger-scale implementation of the program. The new skills can only lead to improved student achievement, the ultimate goal of professional development, if they are applied in practice (Guskey 2002). It would also be interesting to study the effects of the AfL-TPD program on other aspects of student learning, for example, the effects of AfL on student self-regulation.

Acknowledgments

Disclosure statement

Funding

Notes on contributors

Christel Wolterinck is a PhD candidate at the Section of Teacher Development at the University of Twente, the Netherlands. She worked as a chemistry teacher for 14 years. She now is a school leader, working at Marianum in Groenlo, a school for secondary education belonging to the Foundation Carmelcollege. Currently, her work focuses on implementing assessment for learning and data-based decision making via teacher professional development in secondary education and by stimulating research by practitioners in schools.

Cindy Poortman is an associate professor at the University of Twente, the Netherlands. Her research and teaching focus on teacher and school leader professional development in Professional Learning Networks (PLNs). Examples are data teams, teacher design teams, and research practice partnerships. Leadership, sustainability, data use and assessment for learning are her themes of focus. Publications include data use papers and the book The Data Team Procedure (2018, Springer). She also co-edited Networks for Learning (2018, Routledge) and co-edits the Professional Learning Networks Book Series (2019-, Emerald).

Kim Schildkamp is a full professor in the Faculty of Behavioural, Management, and Social Sciences of the University of Twente. Kim's research focuses on data-based decision making and formative assessment. She is a Fulbright scholarship recipient, which she used to study data use in primary and secondary education in Louisiana. She is the previous president of ICSEI (International Congress on School Effectiveness and Improvement). She has published widely on the use of (assessment) data and was, for example, editor of the book "Data-based decision making in education: Challenges and opportunities" and "The data team procedure: A systematic approach to school improvement.

Adrie Visscher is a full professor in the Faculty of Behavioural, Management, and Social Sciences of the University of Twente. In his research he investigates how the provision of various types of feedback to students, teachers and schools (e.g., classroom observation results, students' perceptions of teaching quality, students ability growth, feedback to students who are working on assignments) can support the optimization of the quality of classroom teaching and student learning. As the receipt of such feedback often is the starting point for improvement-oriented actions his research also focuses on the characteristics of effective teacher professionalization.

ORCID

Cindy Poortman http://orcid.org/0000-0001-8133-5985
Adrie Visscher http://orcid.org/0000-0001-8443-9878

Compliance with ethical standards

Ethics committee, Faculty of Behavioural Management, and Social Sciences, University of Twente, approval number: 18,718

References

Baartman, L., and J. Gulikers. 2017. "Doelgericht Professionaliseren : Formatieve Toetspraktijken Met Effect ! Wat DOET de Docent in de Klas ? [Goal-Oriented Professionalisation: Formative Assessment Practices with Effect! What Does the Teacher Do in the Classroom?]" 1–109. https://www.nro.nl/nro-projecten-vinden/?projectid=405-15-722-doelgericht-professionaliseren-formatieve-toetscompetenties-met-effect

Bennett, R. E. 2011. "Formative Assessment: A Critical Review." *Assessment in Education: Principles, Policy & Practice* 18 (1): 5–25. doi:10.1080/0969594X.2010.513678.

Black, P., and D. Wiliam. 1998. "Assessment and Classroom Learning." *Assessment in Education: Principles, Policy & Practice* 5 (1): 7–74. doi:10.1080/0969595980050102.

Carless, D., and K. K. H. Chan. 2017. "Managing Dialogic Use of Exemplars." *Assessment & Evaluation in Higher Education* 42 (6): 930–941. doi:10.1080/02602938.2016.1211246.

Clark, R. E., D. F. Feldon, J. J. G. van Merriënboer, K. A. Yates, and S. Early. 2008. Cognitive Task Analysis. Spector, J. M., Merill, M. D., Van Merriënboer, J. J. G., Driscoll, M. P. eds. *Handbook of Research on Educational Communications and Technology* 3rd (Mahwah, NJ: Lawrence Erlbaum), 577–593.

Cohen, L., L. Manion, and K. Morrison. 2013. *Research Methods in Education*. London: Routledge.

Darling-hammond, L., M. E. Hyler, and M. Gardner. 2017 . ".In *Effective Teacher Professional Development*. Vol. Palo Alto, CA: Learning Policy Institute. doi:10.54300/122.311.

Eggen, T. J. H. M., and P. F. Sanders. 1993. *Psychometrie in de Praktijk*. Arnhem: Cito Instituut Voor Toetsontwikkeling.

Frerejean, J., M. van Geel, T. Keuning, D. Dolmans, J. J. G. van Merriënboer, and A. J. Visscher. 2021. "Ten Steps to 4C/ID: Training Differentiation Skills in a Professional Development Program for Teachers." *Instructional Science* 49 (3): 395–418. doi:10.1007/s11251-021-09540-x.

Frick, T. W., R. Chadha, C. Watson, and E. Zlatkovska. 2010. "Improving Course Evaluations to Improve Instruction and Complex Learning in Higher Education." *Educational Technology Research and Development* 58 (2): 115–136. doi:10.1007/s11423-009-9131-z.

George, P. S. 2005. "A Rationale for Differentiating Instruction in the Regular Classroom." *Theory Into Practice* 44 (3): 185–193. doi:10.1207/s15430421tip4403_2.

Goe, L., C. Bell, and O. Little. 2008. "Approaches to Evaluating Teacher Effectiveness: A Research Synthesis." *National Comprehensive Center for Teacher Quality*.

Gopher, D. 2006. "Emphasis Change as a Training Protocol for High-Demand Tasks." In *AttentionFrom Theory to Practice*, edited by A. Kirlik, A. F. Kramer, and D. A. Wiegman, 209–224. Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780195305722.003.0015.

Guskey, T. R. 2002. "Professional Development and Teacher Change." *Teachers and Teaching: Theory and Practice* 8 (3): 381–391. doi:10.1080/135406002100000512.

Hamodi, C., V. M. López-Pastor, and A. T. López-Pastor. 2017. "If I Experience Formative Assessment Whilst Studying at University, Will I Put It into Practice Later as a Teacher? Formative and Shared Assessment in Initial Teacher Education (ITE)." *European Journal of Teacher Education* 40 (2): 171–190. doi:10.1080/02619768.2017.1281909.

Hattie, J., and H. Timperley. 2007. "The Power of Feedback." *Review of Educational Research* 77 (1): 81–112. doi:10.3102/003465430298487.

Heitink, M. C., F. M. van der Kleij, B. P. Veldkamp, K. Schildkamp, and W. B. Kippers. 2016. "A Systematic Review of Prerequisites for Implementing Assessment for Learning in Classroom Practice." *Educational Research Review* 17: 50–62. doi:10.1016/j.edurev.2015.12.002.

Heritage, M. 2007. "Formative Assessment: What Do Teachers Need to Know and Do?" *Phi Delta Kappan* 89 (2): 140–145. doi:10.1177/003172170708900210.

Kippers, W. B., C. H. D. Wolterinck, K. Schildkamp, C. L. Poortman, and A. J. Visscher. 2018. "Teachers' Views on the Use of Assessment for Learning and Data-Based Decision Making in Classroom Practice." *Teaching and Teacher Education* 75: 199–213. https://doi.org/10.1016/j.tate.2018.06.015

Klenowski, V. 2009. "Assessment for Learning Revisited: An Asia-Pacific Perspective." *Assessment in Education: Principles, Policy and Practice* 16 (3): 263–268. doi:10.1080/09695940903319646.

Kluger, A. N., and A. DeNisi. 1996. "The Effects of Feedback Interventions on Performance: A Historical Review, A Meta-Analysis, and A Preliminary Feedback Intervention Theory." *Psychological Bulletin* 119 (2): 254. doi:10.1037/0033-2909.119.2.254.

Lee, I. 2011. "Bringing Innovation to EFL Writing through a Focus on Assessment for Learning." *Innovation in Language Learning and Teaching* 5 (1): 19–33. doi:10.1080/17501229.2010.502232.

Lynch, R., P. Mannix Mcnamara, and N. Seery. 2012. "Promoting Deep Learning in a Teacher Education Programme through Self-and Peer-Assessment and Feedback." *European Journal of Teacher Education* 35 (2): 179–197. doi:10.1080/02619768.2011.643396.

Mandinach, E., and E. Gummer. 2016. *Data Literacy for Educators: Making It Count in Teacher Preparation and Practice*, Edited by Ed. West. San Francisco, CA: Teachers College Press.

Merrill, M. D. 2012. *First Principles of Instruction*. San Francisco: Wiley.

Petersen-Brown, S., A. R. Lundberg, J. E. Ray, I. N. Dela paz, C. L. Riss, and C. J. Panahon. 2019. "Applying Spaced Practice in the Schools to Teach Math Vocabulary." *Psychology in the Schools* 56 (6): 977–991. doi:10.1002/pits.22248.

Poortman, C. L., and K. Schildkamp. 2012. "Alternative Quality Standards in Qualitative Research?" *Quality and Quantity* 46 (6): 1727–1751. doi:10.1007/s11135-011-9555-5.

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/tea.20163.

Schildkamp, K., F. M. van der Kleij, M. C. Heitink, W. B. Kippers, and B. P. Veldkamp. 2020. "Formative Assessment: A Systematic Review of Critical Teacher Prerequisites for Classroom Practice." *International Journal of Educational Research* 103: 101602. doi:10.1016/j.ijer.2020.101602.

Smith, K. 2016 Cooperative Learning About Assessment for Learning. Laveault, D., Allal, L. eds. Assessment for Learning: Meeting the Challenge of Implementation. 4 (Springer, Cham), 181–197. doi:10.1007/978-3-319-39211-0_11.

Stobart, G. 2008. *Testing Times: The Uses and Abuses of Assessment*. London: Routledge. doi:10.4324/9780203930502.

Vangrieken, K., F. Dochy, E. Raes, and E. Kyndt. 2015. "Teacher Collaboration: A Systematic Review." *Educational Research Review* 15: 17–40. doi:10.1016/j.edurev.2015.04.002.

Van Merriënboer, J. J. G. 2007 Alternate models of instructional design: Holistic design approaches and complex learning. Reiser, R. A., Dempsey, J. V. eds. *Trends and Issues in Instructional Design and Technology*. Upper Saddle River, NJ: Merrill Prentice Hall. 72–81.

Van Merriënboer, J. J. G., and D. Dolmans. 2015 Research on Instructional Design in the Health Sciences: From Taxonomies of Learning to Whole-task Models. Cleland, J., Durning, S. J. eds. *Researching Medical Education*. Chichester, UK: Wiley Blackwell. 193–206.

Van Merriënboer, J. J. G., and L. Kester. 2008. "Whole-Task Models in Education." *Handbook of Research on Educational Communications and Technology* 3: 441–456.

Van Merriënboer, J. J. G., L. Kester, and F. Paas. 2006. "Teaching Complex Rather than Simple Tasks: Balancing Intrinsic and Germane Load to Enhance Transfer of Learning." *Applied Cognitive Psychology* 20 (3): 343–352. doi:10.1002/acp.1250.

Van Merriënboer, J. J. G., and P. A. Kirschner. 2017. *Ten Steps to Complex Learning. Ten Steps to Complex Learning*. London: Routledge. doi:10.4324/9781315113210.

Van Merriënboer, J. J. G., P. A. Kirschner, and L. Kester. 2003. "Taking the Load off a Learner's Mind: Instructional Design for Complex Learning." *Educational Psychologist* 38 (1): 5–13. doi:10.1207/S15326985EP3801_2.

Wiliam, D. 2011. "What Is Assessment for Learning?" *Studies in Educational Evaluation* 37 (1): 3–14. doi:10.1016/j.stueduc.2011.03.001.

Wiliam, D., and S. Leahy. 2015. *Embedding Formative Assessment*. West Palm Beach, FL: Learning Sciences International.

Wolterinck, C. H. D., K. Schildkamp, J. J. G. van Merriënboer, and A. J. Visscher. 2019, January. "Assessment for Learning: A Cognitive Task Analyses in Secondary Education." *Paper presented at the 32nd International Congress for School Effectiveness and Improvement, ICSEI 2019*, Stanvanger, Norway.