Journal of Formative Design in Learning
https://doi.org/10.1007/s41686-026-00115-y

Usefulness and Application of a Co-Designed Formative Assessment Practice Perceived by Teachers in Primary Education

Lars Goertzen[1,2] · **Sylvia Heeneman**[3] · **Trudie Schils**[1] · **Barbara Piskur**[4,5]

Received: 11 March 2025 / Accepted: 16 July 2026
© The Author(s) 2026

Abstract

Formative assessment practices developed in co-design by teachers have the potential to support novel teachers with its implementation. This study examines the perceived usefulness and application of a co-designed formative assessment practice by primary school teachers that were not involved in the co-design process itself. A convergent mixed-method approach was used that integrated both qualitative and quantitative methods. A qualitative descriptive design was used to provide a comprehensive understanding of the program's usefulness as perceived by teachers with or without additional training. A quantitative method was used to evaluate the extent to which formative assessment was applied across both groups. The quantitative data was situated within the qualitative framework for the interpretation of findings. Findings indicate that (i) a co-designed formative assessment practices program supports teachers with the implementation of formative assessment to their local teaching contexts. (ii) A comprehensive adoption, unlike procedural approaches to formative assessment, requires additional support in the form of a complementary teacher professional development program. However, this additional support requires far less time than the time investment put into the co-design process itself.

Keywords Formative Assessment · Co-design · Teacher professional development · Primary education · Mixed-methods design

Introduction

Formative assessment (FA) refers to the elicitation and use of information about pupils' learning to adjust teaching and better meet pupils' needs (Black & Wiliam, 2009). FA strengthens the connection between teaching and learning and is therefore considered a promising approach for improving educational practice (van der Steen et al., 2023). However, teachers often experience difficulties in implementing FA in a coherent and integrated way in daily classroom practice (Veugen, 2023; Yan et al., 2021). Co-design of FA practices in teacher communities is a promising approach to support implementation, as it is situated in practice and produces materials that are considered usable in teachers' own contexts (Goertzen et al., 2023). This raises the question of whether co-designed materials also have practical value for teachers who were not involved in their development — and whether some form of additional professional development support is still needed. This study therefore examines the perceived usefulness and application of a co-designed FA practices program among primary school teachers who were not part of its design. In particular, it compares teachers

✉ Lars Goertzen
l.goertzen@maastrichtuniversity.nl

Sylvia Heeneman
s.heeneman@maastrichtuniversity.nl

Trudie Schils
t.schils@maastrichtuniversity.nl

Barbara Piskur
barbara.piskur@zuyd.nl

1 Department of Macro, International and Labour Economics, School of Business and Economics, Maastricht University, Maastricht, the Netherlands

2 MOVARE School Board, Landgraaf, the Netherlands

3 Department of Pathology, Faculty of Health Medicine and Life Sciences, School of Health Profession Education, Maastricht University, Maastricht, the Netherlands

4 Faculty of Healthcare and Wellbeing, Research Centre for Autonomy and Participation, Zuyd University of Applied Sciences, Heerlen, the Netherlands

5 Faculty of Rehabilitation Sciences, REVAL Rehabilitation Research Center, Hasselt University, Hasselt, Belgium

Published online: 26 August 2026

who received the program without additional support (FA-download group) with teachers who also received training in its application (FA-training group), to better understand what is needed to support adoption across different teaching contexts.

Literature Review

A Comprehensive Approach to FA

FA happens when information about pupils' learning processes is elicited and interpreted to enhance decision-making and to adjust teaching and learning to better meet pupils' needs (Black & Wiliam, 2009). Procedural approaches of FA-like strategies can be relatively simple (DeLuca et al., 2019) and quick to implement, such as the use of mini whiteboards as an all-response system to elicit information about pupils' thinking. However, FA is not complete without an informed follow-up where pupils and teachers have room for adjustment (van der Steen et al., 2022). Consequently, a comprehensive implementation of FA requires extended attention to the integration, coherency, and alignment of different FA strategies in daily classroom practice (Andersson & Palm, 2017; Gotwals & Cisterna, 2022; Wylie & Lyon, 2015). Teachers' implementation of FA is first about teachers' own learning which is why teacher professional development (TPD) is needed (DeLuca et al., 2019; Yan et al., 2021). Explicit attention should be paid to the evolution of teachers' conceptions about FA (DeLuca et al., 2019) (Goertzen et al., 2023) which is why TPD in FA is ideally organized in long-term professional learning programs (Berisha et al., 2024; Thum et al., 2015; Wiliam & Leahy, 2015).

Co-Design as Professional Development

Co-design of FA in a teacher community is a promising approach that supports primary school teachers to design and implement a comprehensive approach to FA (Goertzen et al., 2023). Teacher community co-design is an effective form of professional development that supports teacher learning by being practice-based and fostering active, collaborative participation and shared knowledge construction, while strengthening connections between professional learning and classroom instruction (Desimone, 2009; Ke et al., 2023; Kyza & Agesilaou, 2022; Ouyang et al., 2021). However, to allocate time for teacher community co-design as a form of long-term TPD is a challenge (Zhang et al., 2020). Despite the promise of co-design as an effective way for teachers to implement FA in their daily teaching practice, it is not self-evident that teachers, given the multitude and complexity of tasks, have the required time to participate. However, since co-design leads to usable interventions that are tested by teachers in their teaching practice, the products that are co-designed are based on up-to-date knowledge of good practice, and considered by teachers to be usable in their context (Penuel, 2019).

Practical Value of Co-Designed FA Practices for Adoption by Others

Teacher community co-designed FA practices may have practical value for other teachers in similar educational and socio-economic contexts (Abad & Hattie, 2025). However, the goal of co-design should never be a one-on-one adoption by others since educational methods work only for particular types of learning under particular conditions to reach particular goals (van Merriënboer, 2023). Rather, an adaption to improve contextual fit should be the expectation, depending on how resistant the local teaching contexts are to the newly co-designed product (Ke et al., 2023). Nevertheless, a semi off the shelf educational product may support novel teachers who have limited time to critically assess and refine materials for greater learning impact; however, research on teacher generated materials, particularly collaboratively designed resources and their adoption by teaching peers, remains limited (Abad & Hattie, 2025; Anderson & Shattuck, 2012). For instance, it is unclear how novel teachers who are not involved in the co-design process use a developed educational product and whether some form of additional professional development support is still needed for teachers to implement it in their daily practice. Research on the practical value of co-designed products for novel teachers who are not involved in the co-design process can shed light on the value of co-design regarding adoption by others on a larger scale, and if changes in teaching are sustainable.

Aim

In a previous study, Goertzen et al. (2023) examined how primary school teachers implemented a FA practices program consisting of a set of 14 FA practices that they developed in co-design with researchers and how the teacher-researcher co-design supported the implementation process. This follow-up study's focus is on the perceived usefulness of this FA program by teachers who received it without any further support (FA-download group) and teachers who received additional training in the use of the 14 FA practices (FA-training group). These teachers were not involved in the co-design process as described by (Goertzen et al., 2023). Additionally, the perceived extent to which FA practices are used by both groups (FA-download and FA-training) was investigated to support the paper's focus. A better understanding of the perceived usefulness of the co-designed FA program by teachers and insight into the extent that both

teacher groups use FA is important to design implementation strategies to enhance transfer to local teaching contexts. In this study, we investigate the following research questions (RQs):

(1) How do teachers who downloaded the FA practices program and those who received training perceive the program's usefulness in supporting FA in their classrooms?
(2) Is there any significant difference in the extent to which these two groups use FA practices in their classrooms?

Methodology and Methods

Study Design

To address the RQs, a convergent mixed-methods design was adopted, with qualitative and quantitative data collected almost simultaneously. Given the complexity of educational settings (van Merriënboer, 2023) and the nuances of teacher learning processes (Walkoe & Luna, 2020), a rigorous analytical procedure was used. Quantitative data were embedded within the qualitative strand to account for implementation feasibility that may have influenced teachers' perceived usefulness of the FA program. This integration enabled insights that neither dataset could generate independently and supported a more comprehensive understanding of teachers' perceptions (Skamagki et al., 2024).

The datasets were analyzed separately and integrated at the interpretation stage, with the qualitative strand guiding the analysis and the quantitative strand providing supplementary evidence. Interviews captured teachers' experiences and perceptions of the program's usefulness, while questionnaire data reflected the extent of FA use in daily practice, providing insight into implementation feasibility. Since a program cannot be considered useful if it is not feasible to implement (Randel et al., 2016), this quantitative component helped clarify how actual FA use shaped perceived usefulness and strengthened the link between feasibility and program value.

Semi-structured interviews were conducted near the end of the 2022–2023 school year. The questionnaire (Veugen et al., 2021) was administered in June 2023 via an anonymous online link. Ethical approval was obtained (NVMO-ERB 2022.8.2), and informed consent was collected from all participants.

Sampling and Participants

A combination of criterion-based and convenience sampling was used as a sampling strategy (Moser & Korstjens, 2018) to ensure that participating teachers were affiliated with the same school boards and located in the same region as the schools involved in the teacher–researcher co-design (Goertzen et al. 2023). The study was conducted within a network of two school boards in South Limburg, the Netherlands, covering nearly all primary schools in the region. Schools were invited via email and informed about the study design and participation requirements. Upon consent, teachers could either receive the FA program with a one-page instruction guide (FA-download) or participate in four monthly training sessions regarding the use of the 14 FA practices facilitated by the first author (FA-training). Inclusion criteria were as follows: (a) employment within one of the two school boards, (b) teaching in grade levels 3–8, (c) informed consent, and (d) participation in one of the two conditions.

For RQ1, 34 teachers completed the questionnaire (17 per condition). For RQ2, 23 of these teachers volunteered for semi-structured interviews. Participants represented all grade levels and varied in teaching experience (Table 3, Appendix 1).

The FA-Training Program

The FA-training program was organized solely for teachers that chose to participate in FA-training. The training ran from January to May 2023 and consisted of four sessions. The FA-training sessions were facilitated by the first author with the goal to support the teachers in using the FA practices as a comprehensive unity and facilitate dialogue between the teachers. The co-designed FA practices are presented as activity guides (Goertzen et al., 2023) structured according to two complementary frameworks: the three key processes in teaching, learning, and formative assessment (Wiliam & Thompson, 2008) and the five phases of the FA cycle (Gulikers & Baartman, 2017). For this reason, the content of the first three training sessions was each dedicated to one key strategy, one or two matching phases of the FA cycle, and the accompanying FA activity guides (Table 4, Appendix 2). The final training session focused on combining FA activity guides that originated from the three key strategies (e.g., five phases) to promote FA as a unified practice. Each session lasted 90 min and had the same structure, except for the first session that started with an introduction to FA instead of teachers reporting on their experiences regarding the use of the FA activity guides (Fig. 1). The FA-download group received the 14 FA practices in January 2023, without further support.

Data Collection Instrument

Semi-Structured Individual Interviews

Data was collected using individual semi-structured interviews that captured the perceived usefulness of a

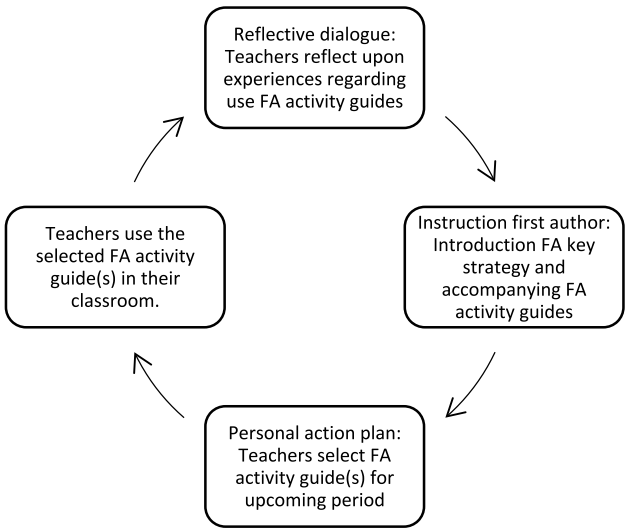

Fig. 1 TPD session structure

co-designed FA program according to novel teachers that were not involved in the co-design process (RQ2). Prior to the interviews, an interview guide was developed (Table 7, Appendix 4). The interview guide touched upon three main topics with several supporting questions: Teachers' motivation for participation and interest in FA, the use of the FA practices program in teachers' daily teaching practice, and teachers' experiences regarding the TPD program (only for FA-training teachers). Peer consultations between the authors of this paper were conducted to get familiarized with the structure of the interview guide, to refine questions if needed, and to promote team-reflexive discussion (Olmos-Vega et al., 2023). The interviews were held and recorded via Microsoft Teams and transcribed by the first author.

Teacher Questionnaire

The teacher questionnaire assessed the extent of FA use in teaching practice 6 months after the FA program (FA-download) or the start of the training (FA-training). Comparing both groups provided insight into implementation feasibility and its relation to perceived program usefulness. The questionnaire comprised 44 items across five subscales representing the phases of the FA cycle (Gulikers & Baartman, 2017; Table 5, Appendix 3). Responses were rated on a 5-point Likert scale from "almost never" (1) to "almost always" (5). It was developed by Veugen et al. (2021), who reported adequate internal consistency (Cronbach's alpha ranging from.72 to.82), discriminant validity (Pearson correlations ranging from.36 to.73), and construct validity based on confirmatory factor analysis. In the present study, internal consistency was also confirmed across subscales (Cronbach's alpha ranging from.77 to.90; Table 6, Appendix 3). The instrument was pilot-tested with non-participating teachers from the same school boards; all items were retained. Participants answered with their mathematics lessons in mind, as mathematics served as the core subject in the co-design study (Goertzen et al., 2023). Data were collected in June 2023 via a secure anonymous Qualtrics survey (Qualtrics, 2022), Version [June 2021] of Qualtrics, Copyright © [2022] Qualtrics.

Data Analysis

Qualitative Analysis

The thematic analysis (Braun & Clarke, 2006) was adopted to analyze the interview transcripts, comprising six-step approach: (1) familiarizing with the data, (2) generating initial codes, (3) generating initial themes, (4) reviewing themes, (5) defining and naming the themes, and (6) writing it up. The first and fourth author independently coded an initial transcript out of each group (FA-download and FA-training) to become familiar with the data and to generate initial codes (step 1). They met to reconcile differences and to create a preliminary codebook. Two more transcripts out of each group were coded by them with the initial codebook. They refined the codes and generated initial themes. An additional debriefing session with all authors was organized to foster mutual understanding and to reach consensus on minor points of disagreement. The first author coded the remaining 17 transcripts (step 2). The first and fourth authors met regularly during the analysis process (debriefing sessions) to discuss the meaning of new codes and potential connections among codes to ensure robust and consistent coding. The first author identified and sorted codes into initial themes and developed a thematic map (step 3). The first and fourth authors reviewed the initial themes by reading all the collated quotes for each theme. They explored whether data within themes was coherent and whether there were clear and identifiable distinctions between themes (step 4). Moreover, they reviewed the validity of individual themes and codes in relation to the data set and renamed some themes and codes to better capture the data (step 5). The first author conducted and wrote a detailed analysis for every theme, identified the "story" of each theme and brought the themes together in a coherent overall story (step 6). The first author wrote a draft and shared it with all authors who reviewed the analytic narrative in relation to the research questions. Since the first author was closely involved in the TPD program and the others only indirectly, the diversity of perspectives enabled authors to shed light on biased assumptions and decision-making (Bieler et al., 2020). Additional insights were discussed and included in the draft. Data analysis software ATLAS.ti 22 was used to support the storage, management, organization, and localization of the research data.

Quantitative Analysis

One aim of this study was to examine between-group differences in the extent to which teachers use FA practices in the classroom, comparing two independent conditions (FA-download vs. FA-training). Given the comparison of two independent groups, a parametric mean comparison test was considered statistically most powerful for detecting group differences, provided that the underlying assumptions were met. We first analyzed the mean total questionnaire scores and subscale scores descriptively (*M*, *SD*) per condition. Secondly, we tested the assumptions of normality and homogeneity of variances using the Shapiro-Wilk test and Levene's test, respectively. The Shapiro-Wilk test showed no significant deviation from normality for either group ($p = 0.224$ and $p = 0.652$, respectively) and homogeneity of variances was confirmed by Levene's test ($p = .451$), indicating equal variances between groups. Finally, given that all statistical assumptions were satisfied, independent sample *t*-tests were conducted to compare the total mean scores and subscale mean scores of the questionnaire between FA-download and FA-training. All teachers who completed the questionnaire provided complete responses. Consequently, no data were missing, and all cases could be included in the analyses. We used the analytic software IBM SPSS.

Reflexivity

The first author has 15 years of experience as a primary school teacher and has extensive knowledge of FA in primary education. He has been involved in the co-design process that led to the FA program. The second author has a biomedical and educational research background and has experience in qualitative research on feedback and assessment systems in health profession education. The third author has methodological experience in quantitative research in an educational context and has been involved in the co-design process that led to the FA practices. The fourth author has methodological expertise in qualitative research but is an outsider in the research context. The second, third, and fourth author had no contact with participants.

We are aware that our personal values, experiences, and beliefs shaped our research and we acknowledge that the analytic process cannot be value-free. From a pragmatic point of view, reality is considered an experience and the truth may be interpreted in terms of the practical effects of what is believed (Savin-Baden & Major, 2023). Therefore, to address our position and influence during this study to balance our own biases, the first and fourth author kept and shared notes of their ideas about the meaning of initial codes during data analysis, and an audit trail was kept of discussions and decisions within the research team. In addition, the diversity of backgrounds, expertise, and perspectives of the authors enabled the authors to shed light on each other's biased assumptions and decision-making (Bieler et al., 2020).

Results

The first part of the results "Teachers' Perceived Usefulness of the FA Practices Program (RQ1)" provides insights in the perceived usefulness of the FA program while the second part of the results "Teachers' Use of FA (RQ2)" explains the extent to which teachers in both groups used FA according to their own perceptions. Finally, we present a working integration model that links the extent in which the teachers use FA (i.e., the FA program's feasibility) and the teachers' perceived usefulness of the program in the.

Teachers' Perceived Usefulness of the FA Practices Program (RQ1)

Three major themes emerged from the analysis in which the usefulness of the FA program was perceived by teachers as "The FA Practices Program Offers Guidance—Is More Than a Toolbox," "Stimulating Teachers' Responsiveness," and "Support and Collective Learning for a Comprehensive Adoption of FA."

The FA Practices Program Offers Guidance—Is More Than a Toolbox

The FA program has been described as easy to integrate, adapt, and align with daily teaching practices. Teachers pointed out that having access to the FA program led to a regular use of FA strategies in their daily teaching practice. The FA activity guides were easy to use because they are user-friendly, clear, and practical.

> Well, I just think the activity guides are very clear. You quickly get an idea of how they work. It is very easy, very concrete to deploy. It is very clear what the intention is, how an activity guide can be used. I can get started with it. It's not hours of extra work. (R11, FA-training)

Teachers mentioned that the FA program encouraged them to give their own twist to an FA activity guide and adapt an FA strategy to their own educational context. Teachers stressed the importance of being able to experiment, reflect, and adapt the scripted FA strategies if necessary, to smooth the progress of their teaching.

> They are just super clear, and the nice thing is that there is also room to do things a little differently if you want to. And that is what the FA activity guides show. You just have to see what suits those pupils you have in front of you. (R7, FA-training)

Teachers explained that another reason for using the FA program lies in the recognizability of the FA activity guides that confirmed their experience-based knowledge. Teachers recognized some scripted FA strategies as being part of their existing teaching practice (i.e., instructional model) which led to the awareness that these strategies had been used without conscious intention during the lesson until that moment. The awareness that some FA strategies were already part of their existing practice stimulated teachers to use these strategies purposefully, and to explore how they could link new activity guides to their existing teaching practice.

> I can see the advantages of combining it with EDI [instructional model], because actually it's very much an extension of that, it just overlaps very nicely. And then you can integrate it that way. (R15, FA-download)

For teachers, the FA program is more than just a toolbox; it enabled them to align FA strategies with existing teaching practices, to make more conscious choices during the lesson and to use as a constructive resource book offering guidance and direction.

Stimulating Teachers' Responsiveness

The FA program supported teachers with eliciting evidence of pupil learning which informed intentional feedback to support pupil learning. The FA activity guides enabled teachers to respond more actively to relevant pupil learning needs and at the same time providing effective feedback and plan follow-up actions within and across lessons. Teachers reported that the FA activity guides that focus on eliciting evidence of pupil learning supported them and enhanced planning the next step in learning. For the teachers in the FA-download group, this was mostly the next step within a lesson. For instance, using mini whiteboards to check for pupil understanding after instruction. The teachers in the FA-training group mostly used the elicited evidence for decision-making to plan follow-up actions across lessons. The FA activity guides that focus on eliciting evidence of pupil learning promoted planning series of lessons instead of standalone lessons for each day. This meant that these teachers put less emphasis on the prescribed program of their instructional method and were guided by the current learning needs of their pupils at that moment. The teachers planned checkpoints and used FA activity guides to gather evidence about their pupils' learning to inform and prepare follow-up actions. Preparing follow-ups mostly meant that teachers would decide whether to go ahead as planned the next day or to delay the curriculum to create space to re-instruct (part of) the group.

> Working, not from the teaching method, but from the learning trajectories, from the data we have, we are going to look, how do we proceed? How are we going to shape the next period? (R1, FA-training)

Teachers mentioned that the FA program improved the use of intentional feedback to support pupil learning. For the FA-download group, feedback was at the level of providing direct verbal feedback on tasks during the lesson. For the FA-training group, feedback was at the level of actionable feedback aimed at helping the pupils move forward in their learning. According to the FA-training group, feedback is most effective when it is related to the pre-established success criteria of a given learning objective. These teachers matched their feedback to the pre-established success criteria and provided the pupils just the right amount of information to enhance insight in their own learning.

> To completely dissect an assignment and then to be able to look much better at how it was done. To pupils who then went to check their work, I said, what step are you at? (pupil) Oh yes, I'm at this success criterion. They get a lot more insight into this themselves. (R2, FA-training)

Support and Collective Learning for a Comprehensive Adoption of FA

To adopt a comprehensive adoption of FA that moves learning forward, teachers need support from a knowledgeable other, should be able to learn collectively, and should focus on connecting newly acquired knowledge of FA with existing knowledge of classroom practice. Although teachers reported that the FA program is useful for improving teachers' professional development, it was also perceived that professional learning in FA needs support, i.e., in the format of a professional development program. Firstly, it was felt that a knowledgeable facilitator is needed because teachers need to have a certain theoretical background to fully apply FA. Although the FA activity guides look like they can be used right away, in-depth knowledge of FA is needed to properly apply them in daily teaching practice, which can be provided by a knowledgeable facilitator. Teachers of both groups mentioned that without support from a knowledgeable facilitator, the use of the FA activity guides may be limiting to performing "tricks" and standalone activities, resulting in a fragmented adoption instead of a comprehensive adoption that improves teachers' instructional practices. The

knowledgeable facilitator can be a colleague within school that is trained and can support other teachers using the FA activity guides.

> You get that information during training. Then you start discussing it and get to work. Whereas if you only give me those activity guides, I'll do a bit of searching myself, and then you just miss the background theory. Now [with training] you try something out with that knowledge. While otherwise it is a bit of trying at random. (R7, FA-training)

Secondly, the importance of a supporting professional development program lies in the opportunity for teachers to learn collectively and to connect newly acquired knowledge about FA with existing knowledge of classroom practices. Teachers should be able to share and adopt ideas, discuss practical difficulties when using specific FA activity guides, prepare and plan FA activity guides together, and reflect together on previous experiences.

> Sessions where you work together with colleagues to prepare things that you can use in the classroom. And then feedback in between, what did we do, how did it go? Exchange examples with each other? (R17, FA-download)

Teachers' Use of FA (RQ2)

Table 1 shows that the FA-download teachers had a mean perception of 3.50 (SD = .48) and the FA-training teachers a mean perception of 3.57 (SD = .60) for their total FA cycle practices. The FA-download teachers used FA practices regarding *clarifying expectations* the most according to their own perceptions (Phase 1; M = 3.73, SD = .52). The FA-training teachers perceived FA practices regarding *eliciting pupil responses* (Phase 2; M = 3.77, SD = .76) as being used the most. The least used FA practices according to the FA-download teachers' perceptions were those associated with *taking follow-up actions: adapting teaching and learning* (Phase 5; M = 3.19, SD = .58). The FA-training teachers perceived *communicating with pupils about responses* (Phase 4; M = 3.30, SD = .84) as the least used FA practice.

Table 2 Independent sample *t*-test results for questionnaire total scores and subscale scores

	t-value	*p*-value (one-sided)
Total	−.41	.34
Phase 1	−.11	.46
Phase 2	−1.34	.10
Phase 3	−.37	.36
Phase 4	.39	.35
Phase 5	−1.37	.09

An independent sample *t*-test was conducted to assess the difference between the perceived use of FA practices of the FA-download teachers and the FA-training teachers. We tested the null hypothesis of no difference in perceived use of FA between the two groups, using a one-tailed *t*-test in which we assumed that TPD in FA leads to more extensive use of FA practices by the FA-training teachers. However, the findings suggest that the perceived use of FA practices was comparable (Table 2).

Linkages Between the Qualitative and Quantitative Findings

When situating the quantitative data (i.e., descriptive statistics and results of the independent sample *t*-test) within the qualitative framework (i.e., codes and themes), several linkages can be established that either converge or diverge with the qualitative findings (Fig. 2). Comparability between the teachers' perceived use of FA in general (QUAN) converges with the teachers' notion of conceptual similarities between FA and the commonly used model of direct instruction (QUAL) as well as with the teachers' report that the FA practices program was easy to integrate in daily teaching practice (QUAL). *Eliciting pupil responses* was perceived as being used the most by the FA-training teachers (QUAN) and converges with teachers reporting that FA activity guides that focus on eliciting evidence of pupil learning supported them and enhanced planning the next step in learning (QUAL). *Communicating with pupils about responses* was perceived as being used the least by the FA-training teachers (QUAN) and diverges from the teachers' notion that the FA practices program improved the use of intentional feedback (QUAL). *Taking follow-up actions* was perceived as being used the least by the FA-download teachers (QUAN) and diverges

Table 1 Descriptive statistics of questionnaire scores total and five phases based on Gulikers and Baartman (2017), for FA-download and FA-training

	Total		Phase 1		Phase 2		Phase 3		Phase 4		Phase 5	
	Download	Training	Download	Training	Download	Training	Download	Training	Download	Training	Download	Training
n	17	17	17	16	17	15	16	15	16	15	17	17
M	3.50	3.57	3.73	3.75	3.44	3.77	3.55	3.62	3.40	3.30	3.19	3.49
SD	.48	.60	.52	.55	.66	.76	.56	.56	.57	.84	.58	.64

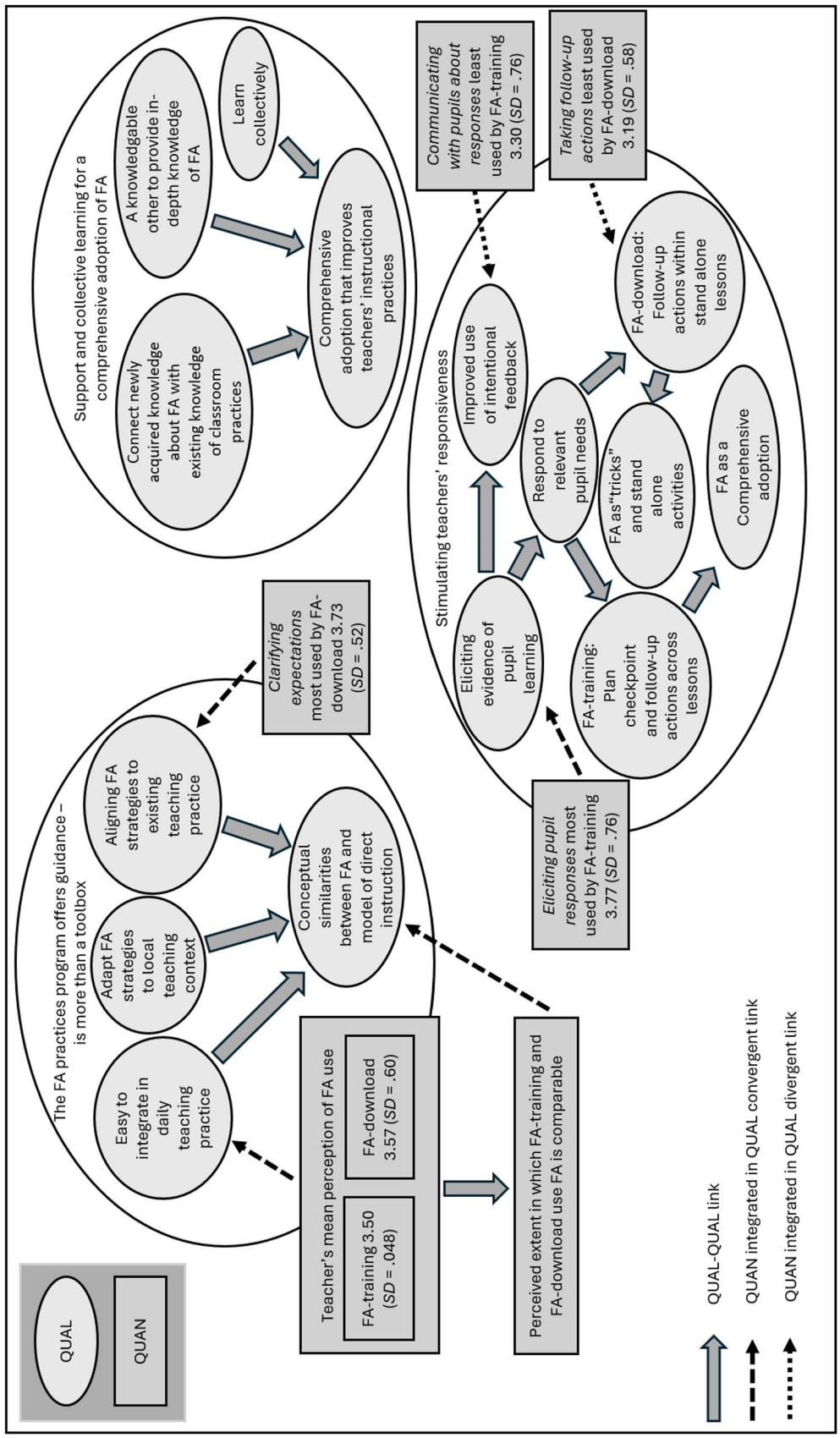

Fig. 2 Working integration model: Integrating the extent of the teachers' use of FA practices (QUAN) in the teachers' perceived usefulness of the program (QUAL)

from the report of these teachers that the FA practices program improved informed decision-making for follow-up actions within the lesson.

Discussion

The main aim of this study was to investigate the perceived usefulness and application of a teacher co-designed FA program by teachers with and without additional support to gain a better understanding of how a co-designed FA program can accommodate novel teachers to implement FA. A convergent mixed method approach was adopted to integrate the quantitative findings within the qualitative framework. This resulted in a working integration model consisting of convergent and divergent linkages between qualitative and quantitative findings. The findings of the one-tailed *t*-test suggesting that the perceived use of FA practices is comparable for both groups fits well with the teachers' reports that the program is easy to integrate, adapt, and align with daily teaching practices. The findings of the *t*-test also fit well with the teachers' notion of conceptual similarities between FA and the teachers' instructional practices. Descriptive statistics however did suggest differences in the extent to which both teacher groups perceive the use of FA strategies. These differences can be illustrated by convergent and divergent linkages between the qualitative and quantitative findings as shown in the working integration model. In this section, we will discuss and substantiate the importance of the integrated findings.

The FA program was indicated as useful to accommodate the procedural and fragmented implementation of FA as well as a more comprehensive implementation of FA (DeLuca et al., 2019). Teachers confirmed that the easy-to-use FA program offered guidance and direction to their teaching, and improved their responsiveness to relevant pupil learning needs which made it more than just an instructional toolbox. However, there were differences in the way that teachers' responsiveness improved as identified in the integrated data (Fig. 2), indicating that both teacher groups differed in the way that they conceptualized FA over time (Lysaght & O'Leary, 2013). Our qualitative analysis suggested that teachers in the FA-training group adopted a more in-depth conceptual understanding of FA that changed the way they think about their instructional practice and their lesson preparations. These teachers planned the FA process as a series of connected formative activities (Furtak et al., 2016) within series of lessons. An example of connecting FA strategies is relating feedback to the pre-established success criteria of a given learning objective (Goertzen et al., 2023). However, the integrated findings show that the reported improved use of intentional feedback diverged with FA-training teachers' use of the FA strategy communicating with pupils about responses. These teachers perceived this FA strategy as being used the least (Fig. 2). A similar finding was reported by Brooks et al. (2019) in their study on the most common types of feedback in upper primary education. They conclude that feed forward (i.e., information used by the learner for improvement) was the least frequent feedback type observed in the classroom. On the other hand, the conceptual understanding of the teachers of the FA-download group appeared to be more procedurally orientated. They conceptualized FA as a collection of fragmented and stand-alone strategies which Marshall and Drummond (2006) identify as the "letter" of FA. Descriptive statistics supported this finding as the FA-download teachers perceived the FA strategy taking follow-up actions: adapting teaching and learning as being used the least. A similar finding was reported by Veugen et al. (2021), in their study on teachers' use of FA activities in the classroom. The participating teachers in this study perceived this strategy also as being least used. FA is not complete without an informed follow-up where teachers have room to adjust their teaching and pupils have room for improvement (van der Steen et al., 2022). Contemporary views on FA include the notion that FA is a planned and ongoing process (Wright & Miller, 2018), consisting of a unified practice of integrated strategies (Andersson & Palm, 2017) that are embedded within and essential to teaching and learning (Gotwals & Cisterna, 2022). Marshall and Drummond (2006) identify this approach to FA as the "spirit." To shift from the letter to the spirit requires teachers to cross a conceptual threshold (DeLuca et al., 2019) because adopting the spirit of FA is not about mastering FA techniques that can easily be implemented into everyday pedagogical practice (Pedder & James, 2012). In contrast to the FA-download teachers, the FA-training teachers reported that they planned checkpoints and follow-up actions across lessons. This indicates that they adopted an approach to FA that is associated with the spirit of FA, illustrating the practical benefits of the co-designed FA practices program beyond mere superficial and procedural implementation. This aligns with the notion that novice teachers (i.e., FA-download teachers in this study) prefer ready-to-use teaching materials that provide structure and predefined guidance, whereas more experienced teachers (i.e., FA-training teachers in this study) tend to focus on achieving deeper pedagogical impact (Abad & Hattie, 2025).

Despite differences in the descriptive statistics on teachers' perceived use of FA strategies, these differences did not reach statistical significance, indicating that the FA activity guides are similarly useful with or without additional support for teachers. As illustrated in the working integration model (Fig. 2), comparable use in both groups may be not only due to the fact the FA program was easy to integrate, adapt, and align with daily teaching practices, but also because of the conceptual similarities between FA and the day-to-day instructional practices that were already being used by teachers. The model of direct instruction, that is most commonly used in Dutch schools (Inspectie van het Onderwijs, 2021), contains formative elements such as clarifying the learning objective, checking for understanding by using mini whiteboards, and providing feedback (Goertzen et al., 2023). In addition, FA is increasingly associated with theories

of pedagogy, instruction, and learning (Black & Wiliam, 2018). FA strategies are therefore not always equally distinguishable from the lesson components of instructional models that are supported with theories of pedagogy, instruction, and learning. This resonates with our integrated findings that some FA strategies were implicitly already part of teachers' teaching practice. The FA-download teachers perceived the FA strategy *clarifying expectations* as being used the most. Explaining the learning objective of a lesson to pupils is considered a common starting point in primary education in the Netherlands (Inspectie van het Onderwijs, 2022) which makes it difficult to determine whether the teachers' perceived use of *clarifying expectations* can be attributed to using the FA program or to the teachers' existing instructional practices. Another reason for comparability in FA use across both groups may be attributed to the level of conceptional understanding of FA as discussed earlier in this section. Teachers' underlying beliefs that support their pedagogical choices (Heitink et al., 2016) and conceptual understanding of FA (Lysaght & O'Leary, 2013) play an important role when implementing a comprehensive adoption of FA (i.e., the spirit of FA). A superficial and procedural adoption of FA will most likely not influence instructional practices fundamentally (Marshall & Jane Drummond, 2006) which is why FA-download teachers' reports on their FA practices may not have been so different from their instructional practices as they were before receiving the FA activity guides. Nevertheless, based solely on our quantitative analysis, the existing FA program, co-designed by teachers in a long-term TPD program, seems to support teachers in their limited time to implement FA, and led to a similar perceived feasibility of implementation in both groups.

Although co-designing FA practices can help teachers to cross the conceptual "spirit to letter" threshold (Goertzen et al., 2023), the present study suggests that the co-designed product is not sufficient in doing so on its own and still needs to be accompanied with teacher professional support in the form of a TPD program with a knowledgeable facilitator that leads the process (DeLuca et al., 2019). Teachers reported that a knowledgeable other (i.e., field expert, lead teacher) can provide a theoretical background that is necessary to fully apply FA. A deep understanding of FA serves as the foundation on which teachers base decisions about their daily teaching practice (Sims et al., 2025). However, teacher professional learning can be a time-consuming and complex endeavor (Desimone, 2009) and may create tensions with teachers' work schedule (Collin & Smith, 2021; Randel et al., 2016). This also applies to co-design between teachers and researchers as a framework that aims at developing integrated professional knowledge (Warr & Mishra, 2021), supporting the development of implementation (Penuel et al., 2022), and providing teacher support to foster the design process (Pieters et al., 2019). An important finding of our study is that additional support for teachers regarding the use of this co-designed FA program does not have to be a time-consuming process that puts a big strain on teachers' workload, as four 90-min training sessions already led to positive outcomes in the way teachers conceptualize and implement FA. This suggests that an FA program, co-designed and tested by teachers, can effectively support novel teachers in a TPD program to implement FA, while remaining feasible and manageable within their existing workload. By focusing on this underexplored area, this study contributes new insights into how FA materials, co-designed by teachers, can accommodate the implementation of FA by peers. Importantly, this study also provides insight into the level of professional support needed for teachers to change their FA practices, which has not been researched before (Veugen et al., 2024).

Practical Implications

The study offers suggestions for future TPD programs aimed at implementing a comprehensive adoption of FA. It highlights the potential of a teacher-researcher co-designed FA program as a "semi-off-the-shelf" educational product to accommodate novel teachers in their limited time to implement the complex educational concept that FA is. Although the FA program was tested by the teachers during the co-design process and was based on up-to-date knowledge of good practice, one-on-one adoption by others should not be expected and strived for. A TPD program needs to provide novel teachers with sufficient time and room to reflect upon and adapt FA practices to empower teachers to make educational changes that are meaningful for their local context (Ke et al., 2023) and improve contextual fit (Penuel, 2019) based on their classroom expertise (Kyza & Agesilaou, 2022). Moreover, as teachers' conception of FA eventually determines the way in which FA is implemented (Lysaght & O'Leary, 2013), it is important to pay explicit attention to teachers' concept evolution to ensure that a comprehensive adoption to FA is being pursued by teachers and to set the pace of TPD process (Goertzen et al., 2023). Finally, as stated earlier, a field expert or lead teacher with advanced theoretical knowledge of FA is needed to lead the TPD program considering that FA is a multifaceted concept.

Limitations and Implications for Future Research

We are aware that our study is small scale, which may limit the generalizability of our findings. In addition, convenience sampling was part of our sampling strategy, with teachers from specific school boards volunteering, which may have induced bias. However, the same approach was applied to both FA conditions, and for RQ2, no condition-specific selection mechanism was introduced, ensuring a fair basis for comparison between conditions. A limitation of the teacher questionnaire is that some FA subscale items may also have touched upon elements of direct instruction, which is part of most teachers' existing

practices. As FA and direct instruction cannot be clearly separated conceptually, the questionnaire data may be clouded. In addition, the questionnaire focused on teachers' practices rather than their conceptual understanding of FA. Future research could include questions on teachers' conceptual understanding to better examine how FA is used in practice and whether teachers adopt a more procedural or comprehensive approach. However, the convergent mixed-method approach used in this study, combining quantitative and qualitative data, enabled us to shed light on the link between teachers' perceived use of FA and their actual approach. Another limitation is that the data were collected anonymously, preventing us from controlling for background characteristics. Such characteristics may influence feedback practices and assessment skills (Christoforidou et al., 2014; Havnes et al., 2012). Finally, data were collected 6 months after the intervention, and it is not self-evident that the effects on teachers' instructional practices will be maintained over time. Future research should therefore examine the longer-term effects of TPD, as professional development is expected to support continuous growth and lasting change (Gaikhorst et al., 2017; Sancar et al., 2021).

Conclusions

FA practices developed in co-design by teachers have the potential to shorten the time investment for teachers outside the co-design to implement FA. This study shows that a co-designed FA practices program supports the implementation of the letter as well as the spirit of FA. Implementing the spirit of FA, however, requires additional support in the form of a TPD program that is supervised by a knowledgeable facilitator. A complementary TPD program consisting of four sessions was a doable time investment for teachers. Thus, a program of FA practices that is co-designed by teachers in practice not only supports other teachers with the transfer of FA practices to their local teaching context but also achieves this in far less time than it took for the development of the initial FA practices program. This makes co-design in education seem more than worth the time investment.

Appendix 1

Table 3 Characteristics of teachers that participated in the semi-structured interviews

		FA-download *n*	FA-training *n*
Total		11	12
Gender	Male	1	2
	Female	10	10
Grade	Grade 3/4[a]	3	2
	Grade 5/6[b]	2	5
	Grade 7/8[c]	6	3
Teacher's years of experience	< 5 years	2	3
	5–15 years	3	4
	> 15 years	6	5

[a]The pupils are approximately 6–7 years of age in Grade 3 and 7–8 years of age in Grade 4

[b]The pupils are approximately 8–9 years of age in Grade 5 and 9–10 years of age in Grade 6

[c]The pupils are approximately 10–11 years of age in Grade 7 and 11–12 years of age in Grade 8

Appendix 2

Table 4 Content overview for each FA-training session

FA-training sessions	FA activity guides	Content
1	1–3	Establishing where learners are going in their learning. Clarifying expectations
2	4–8	Establishing where learners are right now in their learning Eliciting pupil responses Analyzing and interpreting responses
3	9–13	What needs to be done to get where they are going Communicating with pupils about responses Taking follow-up actions, adapting teaching and learning
4	1–13	FA as a unified practice

Appendix 3

Table 5 Teacher Questionnaires with 5-point Likert scale: "almost never" (1), sometimes (2), regularly (3), often (4), "almost always" (5)

1. I formulate learning objectives for the next lesson
2. I formulate learning objectives for a longer period (lesson plans or annual targets)
3. I formulate the criteria for a well-elaborated assignment (success criteria)
4. When I formulate learning objectives, I take the curriculum into account
5. I know the most common misconceptions in my subject/domain
6. I discuss the learning objectives of the lesson with pupils
7. I discuss the learning objectives for a longer period (lesson plans/annual targets) with pupils
8. I discuss the importance of the learning objectives with pupils
9. I check whether the pupils understand the learning objectives
10. I let pupils create their own learning objectives
11. I discuss the criteria for a well-elaborated assignment (success criteria) with pupils
12. I use examples to show pupils what the task should look like
13. I use different FA activities to determine what pupils know and are capable of doing
14. I consciously choose FA activities that fit the learning objectives
15. I use formal FA activities, like pre-testing and interim testing, reports and portfolio discussions
16. I use informal FA activities, like asking deeper questions and observations
17. I use a variety of FA activities to identify the pupils' misconceptions
18. I give pupils the opportunity to show what they have learned
19. I ask pupils to explain their answers
20. I let pupils explain their answers to each other
Based on the information gathered in Phase 2…
21. … I identify the strengths of the whole group
22. … I identify the points for improvement of the whole group
23. … I identify the strengths of individual pupils
24. … I identify the points for improvement of individual pupils
25. … I identify the pupils' misconceptions with regard to the learning topic
26. I take time to analyze the pupils' reactions
27. I compare pupils' work to the criteria for the task
28. I let pupils compare their own work to the criteria for the task
29. I let pupils compare each other's work to the criteria for the task
Based on the analyses of Phase 3…
30. … I discuss the strengths of the whole group
31. … I discuss the points for improvement of the whole group
32. … I discuss the strengths with individuals/small groups
33. … I discuss the points for improvement with individuals/small groups
34. I let pupils think about their own strengths and points for improvement
35. I let pupils think about each other's strengths and points for improvement
36. I give pupils feedback when they are working on an assignment
37. I let pupils improve their work after receiving feedback
38. I provide pupils with tools to give each other feedback
39. I deviate from the learning material/planned targets when I discover that the learning objectives have not yet/already been reached
40. I deviate from the agreed lesson plans when I discover that the learning objectives have not yet/already been reached
41. I help the pupils to come up with the next step in learning
42. I let pupils come up with their next step in learning by themselves
43. I let pupils come up with their next step in learning with each other
44. Based on my previous analyses, I consciously differentiate within my group

Table 6 Cronbach's alpha's of the five scales of the teacher questionnaires

Scale	Teacher questionnaire ($n = 34$)		
	Number of items	Cronbach's alpha	Example Items
Phase 1: Clarifying expectations	12	.84	I discuss the learning goals of the lesson with pupils I know the most common misconceptions in my subject/domain
Phase 2: Eliciting pupil responses	8	.90	I give pupils the opportunity to show what they have learned I use informal FA methods, like asking deeper questions and observations
Phase 3: Analyzing and interpreting responses	9	.85	I compare pupils' work to the criteria for the task I let pupils compare their own work to the criteria for the task
Phase 4: Communicating with pupils about responses	9	.89	Based on the analyses of Phase 3, I discuss the strengths with individuals/small groups I let pupils think about their own strengths and points for improvement
Phase 5: Taking follow-up actions: adapting teaching and learning	6	.77	Based on my previous analyses, I consciously differentiate within my group I help the pupils to come up with the next step in learning

Appendix 4

Table 7 Interview guide

Beforehand

-The interview is held via Microsoft Teams

Introduction (5 min)

-Who am I?

-Why are we here

◦Study on how the FA activity guides are used by teachers

◦This research focuses on the ideas, opinions, and experiences that teachers have regarding the use of FA activity guides

◦How will we proceed

◦Open questions

◦No right or wrong

◦35 min

◦The whole conversation will be recorded via Microsoft TEAMS. The recording will be processed anonymously

-Rules

◦**Phones in silent mode, no other activities on the computer**

Teachers' motivation for using the FA activity guides (10 min)

-Why did you sign up for this project?

-What are your thoughts and ideas about FA?

-What added value does FA bring to your practice?

The use of FA in daily teaching practice (15 min)

-How do you use FA in your daily teaching practice

-How often do you use FA?

-What FA activity guides and combinations of FA activity guides do you prefer?

-Can you tell something about the user-friendliness of the FA activity guides?

-What benefits do the FA activity guides have?

-What disadvantages do the FA activity guides have?

In case the teacher participated in the FA training:

-How did you experience ethe FA training regarding the use of the FA activity guides?

-Are the FA activity guides usable without following the FA training and can you explain why?

Closure (5 min)

-How do we continue from here?

-Are there any questions?

-Thank you

Data Availability The data supporting the findings of this study are available from the corresponding first author upon request.

Declarations

Competing Interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Abad, L. G., & Hattie, J. (2025). The impact of teaching materials on instructional design and teacher development. *Frontiers in Education, 10*,. https://doi.org/10.3389/feduc.2025.1577721

Anderson, T., & Shattuck, J. (2012). Design-based research: A decade of progress in education research? *Educational Researcher, 41*(1), 16–25.

Andersson, C., & Palm, T. (2017). Characteristics of improved formative assessment practice. *Education Inquiry, 8*(2), 104–122. https://doi.org/10.1080/20004508.2016.1275185

Berisha, F., Vula, E., Gisewhite, R., & McDuffie, H. (2024). The effectiveness and challenges implementing a formative assessment professional development program. *Teacher Development, 28*(1), 19–43. https://doi.org/10.1080/13664530.2023.2210533

Bieler, P., Bister, M. D., Hauer, J., Klausner, M., Niewöhner, J. R., Schmid, C., & von Peter, S. (2020). Distributing reflexivity through co-laborative ethnography. *Journal of Contemporary Ethnography*, *50*(1), 77–98. https://doi.org/10.1177/0891241620968271

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability, 21*(1), 5–31.

Black, P., & Wiliam, D. (2018). Classroom assessment and pedagogy. *Assessment in Education: Principles, Policy and Practice, 25*(6), 551–575. https://doi.org/10.1080/0969594X.2018.1441807

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Brooks, C., Carroll, A., Gillies, R. M., & Hattie, J. (2019). A matrix of feedback for learning. *Australian Journal of Teacher Education, 44*(4), 14-32. https://doi.org/10.14221/ajte.2018v44n4.2

Christoforidou, M., Kyriakides, L., Antoniou, P., & Creemers, B. P. M. (2014). Searching for stages of teacher's skills in assessment. *Studies in Educational Evaluation, 40*, 1–11. https://doi.org/10.1016/j.stueduc.2013.11.006

Collin, J., & Smith, E. (2021). *Effective professional development. Guidance Report*. Education Endowment Foundation. https://educationendowmentfoundation.org.uk/education-evidence/guidance-reports/effective-professional-development

DeLuca, C., Chapman-Chin, A., & Klinger, D. A. (2019). Toward a teacher professional learning continuum in assessment for learning. *Educational Assessment, 24*(4), 267–285. https://doi.org/10.1080/10627197.2019.1670056

Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher, 38*(3), 181–199. https://doi.org/10.3102/0013189x08331140

Furtak, E. M., Kiemer, K., Circi, R. K., Swanson, R., de León, V., Morrison, D., & Heredia, S. C. (2016). Teachers'formative assessment abilities and their relationship to student learning: findings from a four-year intervention study. *Instructional Science, 44*(3), 267–291. https://doi.org/10.1007/s11251-016-9371-3

Gaikhorst, L., Beishuizen, J. J. J., Zijlstra, B. J. H., & Volman, M. L. L. (2017). The sustainability of a teacher professional development programme for beginning urban teachers. *Cambridge Journal of Education, 47*(1), 135–154. https://doi.org/10.1080/0305764X.2015.1125449

Goertzen, L., Schils, T., & Heeneman, S. (2023). Co-designing formative assessment practices: A collaboration between elementary school teachers and researchers to conceptualize and implement formative assessment as a unified practice. *Teaching and Teacher Education*, 134, 104306. https://doi.org/10.1016/j.tate.2023.104306

Gotwals, A. W., & Cisterna, D. (2022). Formative assessment practice progressions for teacher preparation: A framework and illustrative case. *Teaching and Teacher Education, 110*,. https://doi.org/10.1016/j.tate.2021.103601

Gulikers, J., & Baartman, L. (2017). Doelgericht professionaliseren: Formatieve toetspraktijken met effect! Wat DOET de docent in de klas? [Targetted professonial development: Formative assessment practices with effect! What the teacher DOES in the classroom]. https://www.nro.nl/sites/nro/files/migrate/Inhoudelijke-eindrapport_NRO-PPO-405-15-722_DEF.pdf

Havnes, A., Smith, K., Dysthe, O., & Ludvigsen, K. (2012). Formative assessment and feedback: Making learning visible. *Studies in Educational Evaluation, 38*(1), 21–27. https://doi.org/10.1016/j.stueduc.2012.04.001

Heitink, M. C., Van der Kleij, F. M., Veldkamp, B. P., Schildkamp, K., & Kippers, W. B. (2016). A systematic review of prerequisites for implementing assessment for learning in classroom practice. *Educational Research Review, 17*, 50–62. https://doi.org/10.1016/j.edurev.2015.12.002

Inspectie van het Onderwijs. (2022). *De staat van het onderwijs 2022*. Inspectie van het Onderwijs. Retrieved from https://www.onderwijsinspectie.nl/binaries/onderwijsinspectie/documenten/rapporten/2022/04/13/de-staat-van-het-onderwijs-2022/IvhO_Staat+vh+Onderwijs+2022_BOEK_def.pdf

Ke, L., Friedrichsen, P., Rawson, R., & Sadler, T. D. (2023). Teacher learning through collaborative curriculum design in the midst of a pandemic: A cultural historical activity theory investigation. *Teaching and Teacher Education, 122*,. https://doi.org/10.1016/j.tate.2022.103957

Kyza, E. A., & Agesilaou, A. (2022). Investigating the processes of teacher and researcher empowerment and learning in co-design settings. *Cognition and Instruction, 40*(1), 100–125. https://doi.org/10.1080/07370008.2021.2010213

Lysaght, Z., & O'Leary, M. (2013). An instrument to audit teachers' use of assessment for learning. *Irish Educational Studies, 32*(2), 217–232. https://doi.org/10.1080/03323315.2013.784636

Marshall, B., & Drummond, M. J. (2006). How teachers engage with assessment for learning: Lessons from the classroom. *Research Papers in Education, 21*(2), 133–149. https://doi.org/10.1080/02671520600615638

Moser, A., & Korstjens, I. (2018). Series: Practical guidance to qualitative research. Part 3: Sampling, data collection and analysis. *European Journal of General Practice, 24*(1), 9–18.

Olmos-Vega, F. M., Stalmeijer, R. E., Varpio, L., & Kahlke, R. (2023). A practical guide to reflexivity in qualitative research: AMEE guide No. 149. *Medical Teacher, 45*(3), 241–251. https://doi.org/10.1080/0142159X.2022.2057287

Ouyang, F., Hu, Y., Zhang, Y., Guo, Y., & Yang, Y. (2021). In-service teachers' knowledge building during face-to-face collaborative learning. *Teaching and Teacher Education, 107*,. https://doi.org/10.1016/j.tate.2021.103479

Pedder, D., & James, M. (2012). Professional learning as a condition for assessment for learning. *Assessment and learning, 3*(2), 33–48.

Penuel, W. R. (2019). Co-design as infrastructuring with attention to power: Building collective capacity for equitable teaching and learning through design-based implementation research. In J. Pieters, J. Voogt, & N. Pareja Roblin (Eds.), *Collaborative curriculum design for sustainable innovation and teacher learning* (pp. 387–401). Springer International Publishing. https://doi.org/10.1007/978-3-030-20062-6_21

Penuel, W. R., Allen, A.-R., Henson, K., Campanella, M., Patton, R., Rademaker, K., Reed, W., Watkins, D., Wingert, K., Reiser, B., & Zivic, A. (2022). Learning practical design knowledge through co-designing storyline science curriculum units. *Cognition and Instruction, 40*(1), 148–170. https://doi.org/10.1080/07370008.2021.2010207

Pieters, J. M., Voogt, J., & Pareja Roblin, N. (2019). *Collaborative curriculum design for sustainable innovation and teacher learning*. Springer. https://doi.org/10.1007/978-3-030-20

Qualtrics. (2022). Qualtrics. In (Version April 2022) *Qualtrics*. https://www.qualtrics.com

Randel, B., Apthorp, H., Beesley, A. D., Clark, T. F., & Wang, X. (2016). Impacts of professional development in classroom assessment on teacher and student outcomes. *Journal of Educational Research, 109*(5), 491–502.

Sancar, R., Atal, D., & Deryakulu, D. (2021). A new framework for teachers' professional development. *Teaching and Teacher Education, 101*,. https://doi.org/10.1016/j.tate.2021.103305

Savin-Baden, M., & Major, C. H. (2023). *Qualitative research the essential guide to theory and practice*. Routledge.

Sims, S., Fletcher-Wood, H., O'Mara-Eves, A., Cottingham, S., Stansfield, C., Goodrich, J., Van Herwegen, J., & Anders, J. (2025). Effective teacher professional development: New theory and a meta-analytic test. *Review of Educational Research, 95*(2), 213–254. https://doi.org/10.3102/00346543231217480

Skamagki, G., Andrew, K., Christine, C., & Wåhlin, C. (2024). The concept of integration in mixed methods research: A step-by-step guide using an example study in physiotherapy. *Physiotherapy Theory and Practice, 40*(2), 197–204. https://doi.org/10.1080/09593985.2022.2120375

Thum, Y., Tarasawa, B., Hegedus, A., Yun, X., & Bowe, B. (2015). *Keeping learning on track: A case-study of formative assessment practice and its impact on learning in Meridian School District*. Northwest Evaluation Association. https://files.eric.ed.gov/fulltext/ED567844.pdf

van der Steen, J., van Schilt-Mol, T., van der Vleuten, C., & Joosten-ten Brinke, D. (2022). Supporting teachers in improving formative decision-making: Design principles for formative assessment plans [original research]. *Frontiers in Education, 7*. https://doi.org/10.3389/feduc.2022.925352

van der Steen, J., van Schilt-Mol, T., van der Vleuten, C., & Joosten-ten Brinke, D. (2023). Designing formative assessment that improves teaching and learning: What can be learned from the design stories of experienced teachers? *Journal of Formative Design in Learning, 7*(2), 182–194. https://doi.org/10.1007/s41686-023-00080-w

Inspectie van het Onderwijs. (2021). *De staat van het onderwijs 2021*. Inspectie van het Onderwijs. Retrieved from https://open.overheid.nl/documenten/ronl-8f43918e-2cfe-4478-b850-9c464f59cfff/pdf

van Merriënboer, J. (2023). *Learning in simulated and real environments*. Maastricht University. https://cris.maastrichtuniversity.nl/en/publications/4df24b33-579b-4f0f-abbc-c34ee6ee405b

Veugen, M. J. (2023). Formative assessment practice in secondary education [internal PhD, WU, Wageningen University]. Wageningen University. https://doi.org/10.18174/635243

Veugen, M. J., Gulikers, J. T. M., & den Brok, P. (2021). We agree on what we see: Teacher and student perceptions of formative assessment practice. *Studies in Educational Evaluation, 70*,. https://doi.org/10.1016/j.stueduc.2021.101027

Veugen, M. J., Gulikers, J. T. M., & den Brok, P. (2024). Secondary school teachers' use of formative assessment practice to create co-regulated learning. *Journal of Formative Design in Learning, 8*(1), 15–32. https://doi.org/10.1007/s41686-024-00089-9

Walkoe, J. D. K., & Luna, M. J. (2020). What we are missing in studies of teacher learning: A call for microgenetic, interactional analyses to examine teacher learning processes. *Journal of the Learning Sciences, 29*(2), 285–307. https://doi.org/10.1080/10508406.2019.1681998

Warr, M., & Mishra, P. (2021). Integrating the discourse on teachers and design: An analysis of ten years of scholarship. *Teaching and Teacher Education, 99*,. https://doi.org/10.1016/j.tate.2020.103274

Wiliam, D., & Leahy, S. (2015). *Embedding formative assessment. Practical techniques for K-12 classrooms*. Learning Sciences International.

Wiliam, D., & Thompson, M. (2008). Integrating assessment with learning: What will it take to make it work? In. Routledge.

Wright, C. M., & Miller, C. M. (2018). Revising the definition of formative assessment. *Council of Chief State School Officers*. https://ccsso.org/resource-library/revising-definition-formative-assessment

Wylie, E. C., & Lyon, C. J. (2015). The fidelity of formative assessment implementation: Issues of breadth and quality. *Assessment in Education: Principles, Policy & Practice, 22*(1), 140–160. https://doi.org/10.1080/0969594X.2014.990416

Yan, Z., Li, Z., Panadero, E., Yang, M., Yang, L., & Lao, H. (2021). A systematic review on factors influencing teachers' intentions and implementations regarding formative assessment. *Assessment in Education: Principles, Policy & Practice, 28*(3), 228–260. https://doi.org/10.1080/0969594X.2021.1884042

Zhang, S., Shi, Q., & Lin, E. (2020). Professional development needs, support, and barriers: TALIS US new and veteran teachers' perspectives. *Professional Development in Education, 46*(3), 440–453. https://doi.org/10.1080/19415257.2019.1614967

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.