In Enhancing Preservice Teachers' Assessment Literacy: Focus on Knowledge Base, Conceptions of Assessment, and Teacher Learning

Päivi Atjonen[1]*, Susanna Pöntinen[2], Sini Kontkanen[2] and Päivi Ruotsalainen[3]

[1] School of Educational Sciences and Psychology, Faculty of Philosophy, University of Eastern Finland, Joensuu, Finland, [2] School of Applied Educational Science and Teacher Education, Faculty of Philosophy, University of Eastern Finland, Joensuu, Finland, [3] UEF Teacher Training School, Faculty of Philosophy, University of Eastern Finland, Joensuu, Finland

Assessment has a critical and time-consuming role in the teaching profession. However, teachers generally possess low levels of assessment knowledge and skills, with their understanding more about the culturally and historically established summative compared to formative assessment. Therefore, developing assessment literacy in preservice teacher education (PsTE) is important. In the present study, a special study module of assessment was designed for PsTE to examine Finnish student teachers' assessment literacy. Deductive and inductive content analyses of students' (N = 168) written diaries were conducted applying a conceptual framework of teacher's assessment literacy in practice (TALiP). Based on the results, the students' assessment knowledge base and assessment conceptions were promisingly versatile and indicated a rich awareness of the broad nature of assessment. However, the key findings also suggest that assessment is not an easy topic for student teachers to discuss with peers. It can be concluded that even short PsTE assessment courses can enhance student teachers' views of versatile pedagogical inputs and outputs of assessment.

Keywords: teacher assessment literacy, assessment and education, assessment conceptions, knowledge base, teacher learning, teacher education – preservice teachers

INTRODUCTION

Teachers spend as much as a third or even half of their professional time involved in assessment-related work (Stiggins, 2014). If the first purpose of assessment is to meet pupils' needs for learning and support, the second and equally important one is to inform teachers about improvements needed in their teaching. As Mellati and Khademi (2018) point out, by employing adequate assessment techniques and grading practices, teachers can improve their instruction, enhance learners' motivation, and increase their levels of achievement.

Do teachers listen enough to assessment data when adapting the pace of instruction, choosing assignments, giving feedback, and deciding on grades, placement, and tracking (Herppich et al., 2018)? Based on international research evidence, the answer seems to be no. Assessment is not taught well enough in preservice teacher education (PsTE) in many countries, and either a lack of advanced courses or effective pedagogy on assessment is evident (Volante and Fazio, 2007; DeLuca et al., 2013, 2019). DeLuca and Bellara (2013) observe that the limited preservice assessment education is theory-laden and disconnected from teachers' daily assessment practices. Similar results were seen also in Atjonen's (2017) research of Finland which is the context of the research reported in this article.

Teachers generally have low levels of assessment knowledge and skills, with beginning teachers particularly underprepared for assessment in schools (Volante and Fazio, 2007; DeLuca et al., 2019). Difficulties with common assessment responsibilities, basic conceptions, and the purposes of assessment are reported (Mellati and Khademi, 2018). Hill et al. (2017) found negative feelings and attitudes toward assessment, and Deneen and Brown (2016) revealed polarized affective conceptions of assessment among preservice teachers. Hill et al. (2017) described how they employ the same assessment strategies they experienced themselves as pupils (see also Hamodi et al., 2016), tend to see assessment as synonymous with the testing culture, and generally abandon formative assessment. To sum, newly qualified teachers' assessment literacy (TAL) level is low.

On the other hand, research findings suggest that when preservice teachers have opportunities to learn about assessment, they demonstrate development (DeLuca et al., 2013). PsTE courses of assessment can be more beneficial by emphasizing formative assessment, avoiding the most traditional assessment methods, and involving preservice teachers as partners in the assessment process (Hill et al., 2017). Assessment-educated PsTE students were more inclined to perceive assessment positively (Yan and Cheng, 2015) and had stronger overall confidence in doing assessment (Charteris and Dargusch, 2018). Positive correlations between self-efficacy and conceptions of assessment are also reported (Levy-Vered and Alhijab, 2015).

In the research reported in the article in hand, we focused on promoting preservice teachers' assessment literacy (TAL). We invited Finnish PsTE students in the research-informed study module to learn about assessment in heterogeneous peer groups. We used the following three pedagogical constructs suggested by DeLuca et al. (2013): perspective-building conversations; praxis: connecting theory to practice, and critical reflection and planning for learning. The special pedagogical viewpoint in the assessment was interaction, i.e., how pupils' participation and educational partnership in assessment processes can be enhanced instead of keeping the summative power solely in the hands of the teacher and school. Our research questions focus on students' assessment knowledge base and conceptions including their collaborative learning about assessment.

From an international research viewpoint, our study enhances the weak tradition of interventions or experiments (2010 onward) that are planned for improving assessment education of PsTE. By our approach, we stand out among the research literature that focuses on the conceptual analysis of TAL (e.g., Herppich et al., 2018; Pastore and Andrade, 2018), is based on student surveys (e.g., Hill et al., 2017), or prefers assessment education standards (e.g., Schneider and Bodensohn, 2017; Wyatt-Smith et al., 2017). To our knowledge, this is a new approach to assessment also in Finnish teacher education.

DEVELOPMENT OF TEACHER ASSESSMENT LITERACY IN PRESERVICE TEACHER EDUCATION

Definitions of Teacher Assessment Literacy

Stiggins (1991) introduced the idea of TAL in his seminal article "Assessment Literacy." Over time, there has been a shift from instrumental and one-dimensional conceptualizations of assessment literacy toward a socio-cultural multidimensional understanding that links to teachers as developing professionals (Willis et al., 2013; Xu and Brown, 2016; DeLuca et al., 2019). This change can be described as a shift from "testing culture" and summative assessment toward "assessment culture" and formative assessment, according to Massey et al. (2020).

Several research examples (e.g., Volante and Fazio, 2007; Stiggins, 2014; Massey et al., 2020) indicate that in-service and preservice teachers perceive assessment mainly as summative tests and grading. This product-oriented approach should be enhanced to include learning processes, i.e., the formative approach needs to be promoted (Cañadas, 2021). These two functions do not exclude each other; assessment literate teachers recognize them both. However, socio-constructivist notions of learning may call for a more rigorous understanding of "assessment for learning" (AfL) (Black and Wiliam, 2018) compared to the culturally and historically established summative approach.

To be assessment literate means primarily having a deep understanding of the interrelatedness of assessment, curriculum, and learning theory, but other advantages are obvious as well. Mellati and Khademi (2018) add that assessment literate teachers can provide understandable information on assessment to educational authorities and parents. Even though parents tend to believe in the ideals of formative assessment (Nieminen et al., 2021), school stakeholders too often rely on testing culture and thus draw inadequate conclusions regarding teaching quality, as Stiggins (2014) points out. Levy-Vered and Alhijab (2015, p. 379) indicate that "assessment-literate teachers are able to draw more valid and reliable inferences about their students' learning and to make better instructional decisions about the content."

Among the important impacts of TAL, confidence in assessment has been explored. Massey et al. (2020, p. 218) explain that "…when individuals feel incompetent and lack confidence in a task, they will choose to engage in it. In contrast, when an individual feels incompetent and lacks confidence in a task, they will avoid engaging in it." Educating assessment-literate educators is also a question of instilling confidence in carrying out assessment, i.e., teachers learn to trust their own skills to diversify

assessment tasks and methods rather than routinely resorting to the most general or typical assessment approaches. Levy-Vered and Alhijab (2015) argue that training in assessment and its multiple conceptions directly affect assessment literacy. Finally, assessment literacy directly and indirectly affects assessment self-efficacy and may result in new assessment practices.

Key Ideas of Teacher's Assessment Literacy in Practice-Model

We prefer the concept "assessment literacy," although DeLuca et al. (2019) mention also the terms "assessment competency" (see e.g., Schneider and Bodensohn, 2017; Herppich et al., 2018) and "assessment capability" (see e.g., Booth et al., 2016; Hill et al., 2017; Charteris and Dargusch, 2018). Our research relies on the model of "teacher assessment literacy in practice" (TALiP-model) designed by Xu and Brown (2016), who use the concept "teacher assessment literacy." Xu and Brown (2016) legitimate their model by reviewing one hundred original research articles on assessment literacy and illustrating it as a pyramid (**Figure 1**).

The TALiP-model includes six components: (A) Knowledge base; (B) Teacher conceptions of assessment; (C) Institutional and socio-cultural contexts; (D) Teacher assessment literacy in practice; (E) Teacher learning; and (F) Teacher identity (re)construction as assessor. Drawing on the socio-cultural approach, Xu and Brown (2016) describe how the knowledge base (elements 1–7) enables teachers to formulate their conceptions of assessment that are interpreted in the national curricular framework, for example. These conceptions have both cognitive and affective dimensions, and they are integrated into the broader views of learning and teaching.

Based on their conceptions, teachers choose different assessment practices (e.g., exams, portfolios, essays, and presentations), which may be successful or may sometimes require a return to the knowledge base or conceptions. Xu and Brown see "assessor identity" as the aim at developing of TAL, which involves teacher learning, i.e., constant reflection on practices and collaboration with colleagues. This long-term process can be successfully launched during PsTE if the developmental scenario is introduced and opportunities to practice some of its components are available.

Based on their praxis-oriented model, Xu and Brown (2016, p. 159) state that "assessment literate teachers are those who constantly reflect on their assessment practice, participate in professional activities concerning assessment in communities, engage in professional conversations about assessment, self-interrogate their conceptions of assessment, and seek for resources to gain a renewed understanding of assessment and their own roles as assessors."

Research Questions

Based on the above-described theoretical framework, we formulated our research questions as follows:

RQ1: How do preservice teachers reflect on the *assessment knowledge base*? (TALiP-component A)

RQ2: What kinds of *conceptions of assessment* do preservice teachers have? (TALiP-component B)

RQ3: How do preservice teachers describe their *collaborative learning* of assessment? (TALiP-component E)

We planned a study module that aimed to increase students' assessment literacy by presenting the key concepts, interactive nature, and methods of assessment (hereafter "assessment module"). After implementing the assessment module, a qualitative evaluation of students' experiences was carried out to answer the RQs.

To our knowledge, this is the first study of the TALiP-model, originally explored and designed for in-service teachers, where it is used to describe preservice teachers' assessment literacy. The TALiP-model did not guide the design of the assessment module because we were not familiar with it while planning the course. Therefore, TALiP is used as the analysis tool for the data that did not include students' texts regarding components C, D, and F.
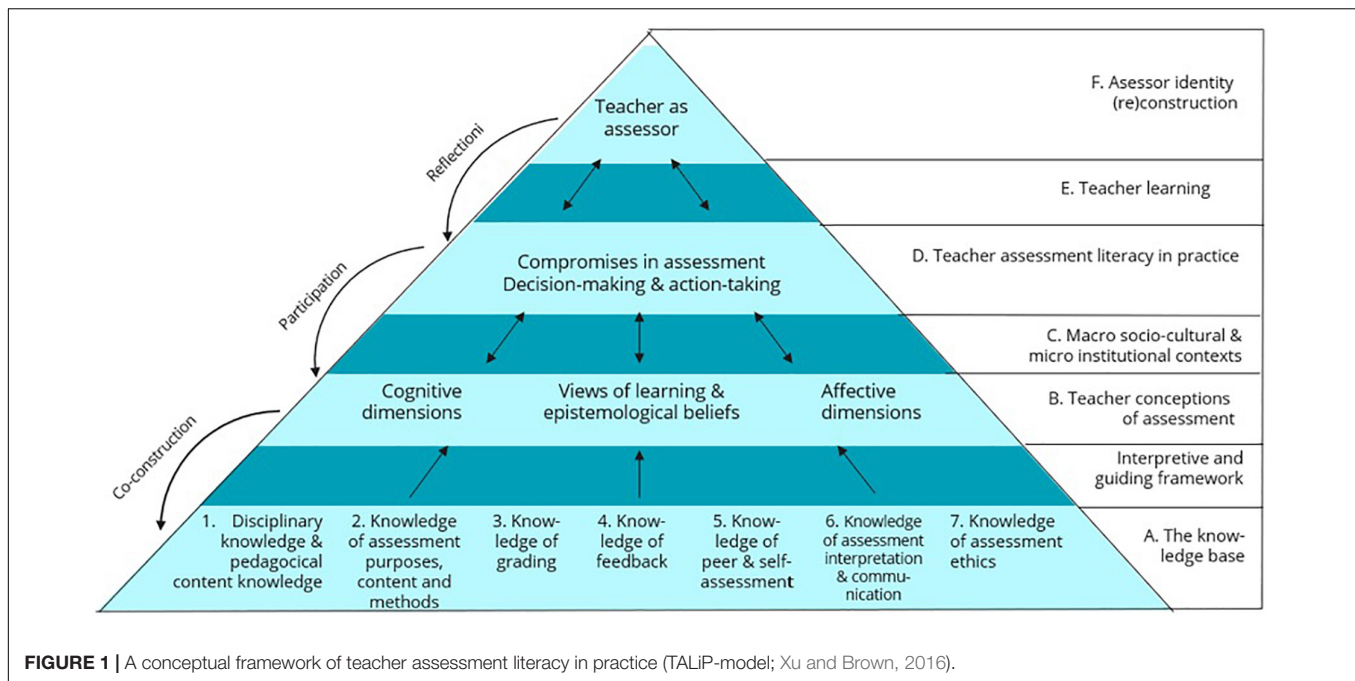
MATERIALS AND METHODS

Finnish Context for Research Results and Their Interpretation

Our assessment module took place in October–December 2020, in the middle of assessment changes in the basic education of Finland (grades 1–9) initiated by the Finnish National Agency for Education (FNAE). In February 2020, renewed guidelines for assessment were published (Finnish National Agency for Education [FNAE], 2020a) to specify the national core curriculum for basic education (Finnish National Agency for Education [FNAE], 2014). At the end of December 2020, new assessment criteria for the final assessment of basic education were introduced (Finnish National Agency for Education [FNAE], 2020b).

As the Finnish educational system is strongly decentralized, municipal educational authorities design local curricula following the national guidelines of the core curriculum. However, despite the acknowledged expertise of teachers and local educational authorities, decentralization can cause unequal assessment practices. Actually, both curriculum-related assessment reforms in 2020 were motivated by rigorous research evidence that assessment was not equal enough in schools all over the country. In addition, new results of the broad national evaluation of assessment practices (Atjonen et al., 2019), conducted during 2017–2019, were also useful for the reforms.

As a country of low-stakes assessments, Finland trusts well-educated teachers' professional knowledge, sees formative assessment as very important, and prefers steering and development over accountability. All teachers in Finland are educated at research-based universities where students take master's degrees, majoring either in education (lower level of basic education) or in their respective subjects (upper level of basic education). Broad pedagogical studies (60 ECTS) are included in their studies.

FIGURE 1 | A conceptual framework of teacher assessment literacy in practice (TALiP-model; Xu and Brown, 2016).

Based on this TE background, Finnish PsTE students should become quite well informed about assessment as an effective pedagogical tool. Because each university decides independently on the PsTE curriculum, there may be a need for special attention to ensure that all prospective Finnish teachers will become provided with basic knowledge on assessment. Therefore, we saw it as important to strengthen assessment literacy among a student cohort of a university in the way that is described next.

Description of the Assessment Module

The second author of this article planned the assessment module together with her teacher educator colleague, and four other experienced colleagues supported them. It tookplace in a TE department in a publicly financed, research-based University of Eastern Finland (UEF) in October–December 2020. The assessment module was the comprehensive part of the theoretical-pedagogical course of "Interaction in learning and educational environments" (ILE; 5 credits) which was compulsory for the whole cohort of students in the academic year 2020–2021. The TE of the UEF has study programs for prospective counselors and teachers of preschool, primary school, secondary school, special education, craft education, home economics, and adult education.

Our study module was not a "traditional assessment course" in which assessment concepts, paradigms, principles, and methods were studied independently of the learning substance. Despite their strengths, in Finnish PsTE, assessment is seldom taught as separate courses. We integrated assessment into the study module that was pedagogically focused on interaction, one of the core concepts of the teaching-studying-learning processes and educational theory. This resonated well with assessment's modern notions of involvement and engagement, and the practical study mode of student peer groups.

Special online learning materials were designed for the ILE, consisting of two modules: assessment theory (e.g., definitions of concepts and methods, school-home cooperation in assessment, and assessment and novice teachers) and assessment digitalization (e.g., interactive elements, formative methods, quizzes, and e-portfolios). Students had to study and discuss the learning material before engaging in their peer group activities. The learning materials presented the basics of summative assessment as well, but we chose to focus more on formative assessment. Students themselves have not often witnessed it during their own school years, and recent Finnish curricular changes emphasize the pedagogical use of formative assessment for learning promotion. Formative assessment fitted well to the underlying theoretical topic of "interaction" in the ILE study module that was available to us to improve assessment literacy.

Five workshops (90 min each) were devoted to planning and modeling a learning object for a digital learning environment, with a special requirement to plan and discuss it from an assessment viewpoint. Students ($N = 382$) were randomly divided into peer groups of 4–5 persons with mixed study program backgrounds. The groups produced a learning object collaboratively, and each student wrote an open diary regarding their personal reflections on the assessment and their discussions in the multidisciplinary peer group.

Regarding research ethics, students knew in advance that the diary was included in the evidence needed for the grading of ILE (university's assessment scale: 1 = poor, 5 = excellent) for each student. We wanted to encourage them to write well-thought diaries which would improve also the quality of our data.

They were also informed that giving or withholding research permission would not have any impact on the course grade. We did not use students' grades in the data analysis.

Data for Empirical Analysis

We explored the written diaries as our empirical data. We asked the students to reply to three diary assignments (DA): (DA1) What did you learn about the assessment? (DA2) What would you like to learn more about the assessment? How would you develop the learning material designed for ILE? (DA3) Did you have shared views and conceptions of assessment in the group, or did they vary? Did you reach a common understanding in your assessment reflections? By numerous thought-provoking questions, we wanted to avoid categorical yes-no-accounts that would have flattened the contents of diaries.

Altogether, 168 students gave permission to use their diaries as research data. The majority were prospective subject teachers (47% of the participants), followed by students of a class teacher education program (26% of the participants). In Finland, class teachers teach all school subjects to pupils of 7–12 years old, and each subject teacher is typically responsible for teaching 2–3 school subjects to 12–15-year-old pupils. The rest, 27%, were studying to become special education teachers, pre-primary teachers, counselors, or adult educators. Regarding the background variables that were available (study program, course grade, and gender), no biases were noticed in the sample (168 students) compared to the population (382 students).

The scope of the fully anonymized data was 33,300 words. The shortest accounts consisted of 45–70 words (10 out of 168 accounts), and the longer ones were approximately 400 words (the longest included 851 words). The middle-sized accounts were typically 100–120 words. The students were coded with randomly chosen numbers, and the DA-number was also included in the citations in the Results (e.g., 2:154 = an account of diary assignment no. 2 written by student no. 154).

One analysis unit was defined as an utterance with a single meaning. Often, the units consisted of a sentence or two.
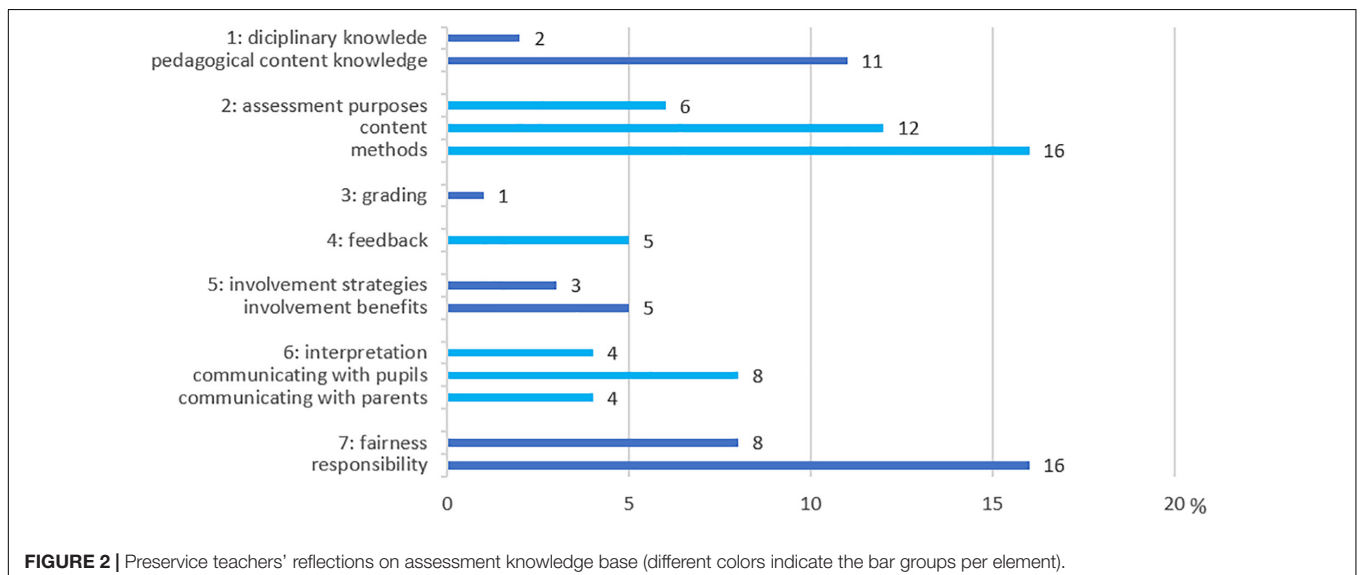
Students' accounts of each DA typically included 2–4 single meanings. Therefore, the sums of the utterances in **Figures 2**, **3** are larger than the number of students. The content analysis was done by means of Excel and ATLAS.ti Web software. Based on shared negotiations regarding the coding protocol and its principles of interpretation, short manuals were written for each DA. The first two authors of the article took on the main responsibility of the analysis.

Analysis of Knowledge Base (RQ1) and Conceptions (RQ2)

To answer RQ1 and RQ2, we addressed the data by deductive content analysis (Neuendorf, 2019; Kyngäs and Kaakinen, 2020), where analysis has prior theoretical knowledge as the starting point and at least a half-structured matrix is used. The data from DA1 in RQ1, and DA1 + DA2 in RQ2 were classified and interpreted from the following viewpoints, relying on the manifest content of the diary texts (Bengtsson, 2016):

- RQ1: Xu and Brown (2016) see the solid *knowledge base* (TALiP-component A) as essential (but not sufficient) for both pre- and in-service teachers to become effective assessment practitioners. The knowledge base is comprised of the seven elements displayed in **Figures 1**, **2**.

- RQ2: Xu and Brown (2016) define *conceptions* (TALiP-component B) to denote the belief systems that teachers have about the nature and purposes of assessment when pupils' learning is examined, tested, evaluated, or assessed. Brown and Gao (2015) classify six inter-correlated assessment purposes (see also Brown et al., 2019). With reference to Xu and Brown (2016), Xu and He (2019) identify six targets of assessment conceptions of preservice teachers. These two categorizations overlap in "purpose," which is one of the targets.

Although the percentages of the distinctive categories of the knowledge base and conceptions are indicated in
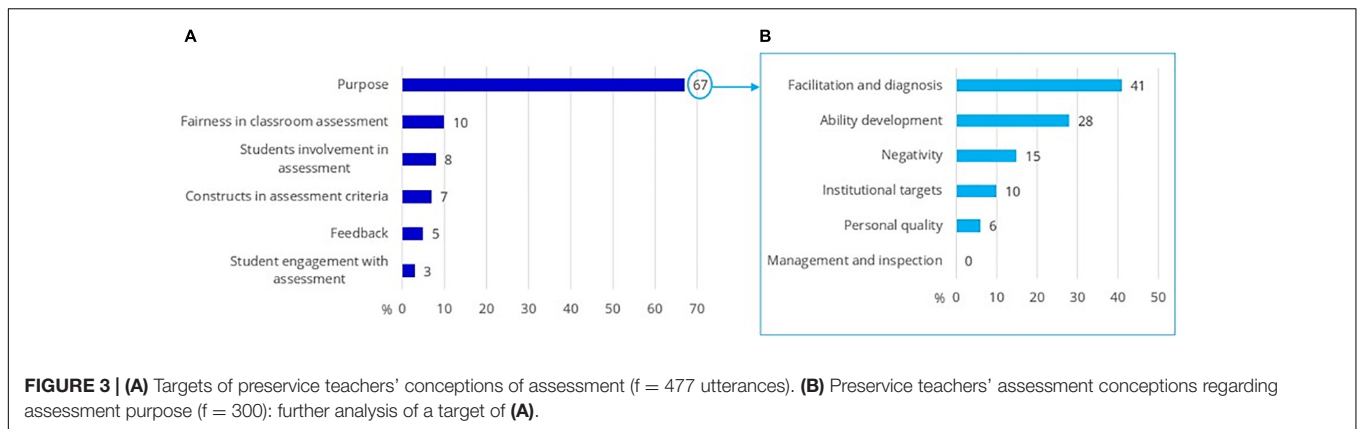

FIGURE 2 | Preservice teachers' reflections on assessment knowledge base (different colors indicate the bar groups per element).

FIGURE 3 | **(A)** Targets of preservice teachers' conceptions of assessment (f = 477 utterances). **(B)** Preservice teachers' assessment conceptions regarding assessment purpose (f = 300): further analysis of a target of **(A)**.

Figures 2, **3**, we do not interpret the results quantitatively. Instead, we use frequencies to facilitate the presentation order of the results and discuss the frequencies as indicative of the heterogeneity in the participants' views. Therefore, more emphasis is placed on the qualitative approach to illustrate preservice teachers' versatile ways of perceiving functions, principles, and methods of assessment.

Analysis of Student Teacher Learning (RQ3)

Student teacher learning (TALiP-component E) was our third focus. To answer RQ3, mainly inductive content analysis was performed on the data of DA3. In the first phase, each of the student teachers' (f = 157) accounts were reduced to keywords that included the main messages. Most preservice teachers described their reflections with words such as *"consensus"* (consensus, congruent, common, together; used 98 times) *or "similar"* (similar, similar-like, parallel, agreed, of the same line, homogenous; used 49 times) meaning that the group members had virtually the same opinion of assessment. These were left out of the analysis because they did not increase any knowledge on student learning. The same was true also with words *sharing* and *differences* (i.e., different, differing, or contradictory) that were rare (used 17 times) and used alone without any supplementary wording.

Second, rest of the keywords (f = 89) were grouped into two main themes. They were discussions and negotiations (f = 51), and compromises and contradictions (f = 38). Third, the themes were qualitatively analyzed which was facilitated by means of two theoretical conceptualizations of collaborative learning published by Nokes-Malach et al. (2015) and Scager et al. (2016).

RESULTS

Knowledge Base of Teacher Assessment Literacy (RQ1)

When all utterances (f = 449) of the assessment knowledge base (KB) were counted, **Figure 2** was constructed to illustrate student teachers' reflections of seven elements based on the TALiP model of Xu and Brown (2016). The students reflected

mostly on assessment methods and responsibility (16 + 16% of the utterances) and least on grading, disciplinary knowledge, and involvement strategies (1 + 2 + 3%).

There are some differences in **Figure 2** compared to **Figure 1** due to the nature of our empirical data regarding preservice teachers (not in-service teachers) and the integrative topic of the assessment module (*interaction* in assessment processes).

- Instead of title "no. 5 Knowledge of peer and self-assessment" in **Figure 1**, the title of "involvement strategies and involvement benefits" is used in **Figure 2** based on the broader title suggested by Xu and Brown (2016) on page 156 in their article.

- Titles no. 6 in both figures vary slightly because we wanted to clarify what "communication" meant in our data.

- "No. 7 Ethics" in **Figure 1** is clarified with the aspects "fairness" and "responsibility" in **Figure 2**.

The first broad finding (not specified to single elements 1–6) from **Figure 2** was the prominence of the pedagogical aspects of assessment in utterances. The students focused on formative assessment, although other assessment purposes were named. They explained, for example, that "in primary education, assessment has two mutually reinforcing purposes, which are formative and summative assessment" (1:70) and that "the familiar methods of assessment are mainly diagnostic, formative, and summative assessment" (1:47).

Assessment purposes and content (KB-element 2) was linked to the core pedagogical benefits of formative assessment: to enhance pupil motivation and achievement (Cauley and McMillan, 2010). The students explored formative assessment as a process that was closely linked to the pupils' learning needs. Strength-based and forward-looking approaches were described as important. As concrete steps toward these ends, they stressed frequent assessment to detect and support pupils' individual skills and capabilities in the long term. For example, students reflected on this knowledge base as follows:

"The purpose of the assessment is to inform pupils and encourage their progress. Assessment should be varied and continuous" (1:64).

"Assessment's purpose it to describe how objectives set for growth and learning have been achieved . . . it is a good tool for a teacher to control studying and learning" (2:47).

The student teachers were committed to applying versatile *assessment methods* (KB-element 2) in their teaching. They combined different assessment methods aimed at giving "pupils the opportunity to demonstrate their skills in a more diverse way" (1:48) and commented that "self-assessment, peer assessment, and assessment work best together as a whole" (1:140). However, students reported that the teacher must choose appropriate assessment methods in determining assessment goals. This was suggested because "the same assessment method does not work for every situation" (1:48) or "it is good to make use of different methods in assessment so that everyone can demonstrate their skills in a variety of ways" (1:17).

Assessment methods, digital assessment included, drew the students' attention, although they did not view digital assessment and traditional assessment separately. Using digital tools in the classroom can help the teacher and pupils meet individual learning needs (Panero and Aldon, 2016). Despite the benefits of digital assessment practices, the students claimed that the use of these tools can reduce interaction, particularly in online learning and distance learning settings. The following examples summarize the students' views that the teacher is responsible for deepening pupils' learning processes and results both in the traditional classroom and in online or distance learning settings:

". . . the digital assessment should also stay with the pupils afterward, so that, for example, after the Kahoot multiple-choice test, we return to discuss the answers" (1:65).

". . . being supportive in the assessment is a complex or even challenging task in face-to-face teaching, online learning will make these even more difficult. As the interaction between teacher and students decreases, formative assessment in particular, can be a major challenge for the teacher" (1:88).

Feedback and pupils' involvement (KB-elements 4 and 5) revealed critical knowledge areas. Finding a balance between positive and developmental feedback was valued by students since "there is something good about every job and it is important to bring it out. Negative feedback should also be brought up as constructive criticism, and it is not good to just give direct negative feedback" (1:123). For example, Li (2019) suggests that pupils must be provided with assessment training before engaging them in complex assessment practices. Student teachers admitted that attention must be paid to the pupils' assessment skills since "assessment also requires pupils to understand feedback and to work on issues where there were shortcomings or to strengthen their own strengths" (1:120).

Communicating with pupils and parents (KB-element 6) illustrated one-way communication. "The teacher states assessment criteria both for pupils and parents, as this kind of activity increases understanding and mutual trust" (1:165), and the teacher's role is "to make assessment more transparent so that both pupils and their parents know what is being assessed and how" (1:73). Obviously, knowledge concerning mutual communication can be tricky for the students to reflect on since their experiences of teaching and home-school cooperation might be limited.

The reflections involved in interpreting the assessment data to deepen professional knowledge remained vague but promising. As Timperley (2009) suggests, assessment data is used not only to make the links between particular teaching activities and what the pupils actually learn but also to guide teachers to change their teaching practices. The students clearly noticed that assessment data "helps identify gaps in teaching and improves one's own performance, as through assessment, you get the guidelines" (1:47) and "the role of assessment remains minimal if the teacher is not capable of directing her or his teaching with the "assessment information" she or he collected" (1:53).

Since knowledge of *assessment ethics* (KB-element 7) gives teachers the opportunity to engage in assessment at a deeper level (Xu and Brown, 2016), it was good to notice that ethics was reflected in the fairness and responsibility perspectives. This represents that a general body of student teachers' understanding of assessment included statements about the fairness, equity, and social justice of assessment. *Responsibility,* in turn, referred to clarifying the complex nature of assessment and of implementing assessment practices. In other words, re-thinking the nature of assessment and attempts on how to improve assessment practices in everyday settings seemed to conceptualize the ethical points of assessment for the prospective teachers. The following examples summarize well the preservice teachers' attempts to engage in implementing transparent, fair, and accurate assessment practices:

"I learned that in addition to development targets, assessment should pay attention to strengths, however, in a way that assessment is fair to everyone" (1:17).

"Above all, I learned how complex the assessment is and what aspects a good assessment should consider. . . . One of the issues raised was how difficult it can sometimes be to make a fair assessment, as the teacher may lack relevant information. Therefore, in conducting the assessment, you must always be very accurate and careful" (1:156).

Conceptions of Assessment (RQ2)

The TALiP-model indicates that teachers' versatile conceptions of assessment are distilled from the knowledge base mainly through the curriculum as the guiding framework. Based on the deductive content analysis of DA1 and DA2, utterances ($f = 447$) of assessment conceptions were condensed in **Figures 3A,B** to illustrate student teachers' reflections.

As **Figure 3A** indicates, most utterances regarding assessment conceptions focused on assessment purposes. Regarding various assessment purposes illustrated in **Figure 3B**, the students saw that assessment's main purpose was to *recognize and support pupils' needs* ("facilitation and diagnosis") and to *promote pupils' ability development*. Based on the research literature, these two

are undoubtedly the most important pedagogical purposes of assessment. Therefore, it is encouraging to notice that the students had understood this already in their PsTE phase. For example, the students described their conceptions as follows:

> "Assessment is a whole that has enormous influence on a pupil's learning and self-concept" (1:11).

> "Assessment's purpose is to describe how the objectives set for growth and learning have been achieved … it is a good tool for a teacher to manage studying and learning" (2:47).

> "In assessment, pupils must be noticed as individuals regarding their development and skills levels" (1:55).

> "The purpose of assessment is to inform students about progress and to encourage them" (1:64).

Reflections on the *institutional or managerial* purposes of assessment were not numerous, but conceptions of assessment's *negative functions*, bad reputation, or critical consequences are worth noting here. On the one hand, they mentioned negative aspects, such as "assessment is sometimes very challenging for both pupils and teachers" (1:147) or "many pupils are overloaded because of continuous assessment, including an increased number of tests and projects" (2:139). On the other hand, they talked about how to cope with pupils' negative experiences of assessment:

> "Assessment has the danger of giving pupils a feeling of failure … therefore, it is important to explain to pupils about assessment methods so that they understand the grounds of their grades" (1:133).

> "… try to avoid starting with negative feedback, but say first what was successful and then focus on criticism" (2:105).

> "Pupils' trust in assessment may be diminished by deficient information of assessment criteria" (1:80).

Approximately a third of the conceptions dealt with other targets than the purpose (**Figure 3A**). Many students expressed two ethically topical and inter-related assessment conceptions: they wrote diligently about the teacher's task to care for the *fairness in classroom assessment* or used explicitly the word "(assessment) criteria" included in the category of *constructs in assessment criteria*. The students described fairness with adjectives such as fair, unbiased, honest, truthful, transparent, and valid. They discussed fairness in relation to the versatile use of methods in enabling the visibility of pupils' learning (e.g., 1:10, 1:17, 1:96) and emphasized teachers' cooperation in the assessment of the same pupils (e.g., 2:4, 2:35). They referred also to biases in relation to careless documentation (2:34), ignorance of individual differences (1:55), and negligence of developmental orientation (1:60).

The difference between the two conceptions of *pupils' involvement* and *engagement* (**Figure 3A**) was apparent: the student teachers were more concerned with the equal distribution of power between teacher and learners and felt that the learners' voices should be better heard in assessment. They explained how "it is important to let pupils to participate in assessment more often" (1:2) or "a pupil should decide together with the teacher

what is the best way to show what is learned" (1:6). Less was written about how to succeed in engaging pupils themselves to improve their learning by means of assessment ["pupils should learn something from assessment and, based on that, try to develop their activities in the future" (1:43)].

Student Teachers' Collaborative Learning (RQ3)

Near the top of the TALiP pyramid is the "teacher learning" (component E), in our case, "student teacher learning." Xu and Brown (2016) argue that assessor identity construction is possible by means of frequent reflections and participation in learning communities with colleagues, in our case, with "peer students." The DA3 enquired about peer group reflections.

The key result was *unanimity* as was anticipated in the analysis section above. Discussions were seen as "superficial" (3:144), "tentative" (3:35), or "insufficient" (3:42). Probably, some students were not provided with proper basic knowledge on assessment (concepts, methods, and principles), did not clearly recognize the assessment element of this ILE study module, or were too much stuck on their own school experiences of assessment (e.g., 1:15, 2:37, 3:4, 3:35).

Regarding the two main topics in the second phase of analysis, students reflected upon group *discussions or negotiations* (words as discuss, negotiate, ponder; $f = 51$) and recalled contributions with *versatile, even contradictory viewpoints ($f = 38$)*. Several students indicated the added value of cooperation and talked about *observational learning, increased engagement, or negotiating multiple perspectives* (Nokes-Malach et al., 2015):

> "In our discussions, we complemented each other very much. A team member started to talk, and another member enhanced it but also dared to bring in even very different opinions" (3:60).

> "With the help of discussions, we were able to enhance our own personal views of assessment" (3:88).

> "It was particularly important to have common critical discussion so that everybody knew others' viewpoints and their rationale" (3:97).

Longer accounts in the DA3 data were valuable because they illuminated constructions of shared thoughts and struggles for consensus. The following two excerpts exemplify how the two groups tried to enrich the discussions and reach a common understanding of assessment and its use in learning.

> "Assessment … caused significant reflections due to differences of opinion. These differences focused on formative assessment and whether it should be included or excluded. I saw the final compromise as reasonable and thanked Mary [pseudonym] for her perceptive comments. She did not look grumpy, and she accepted it. I can say that differences of opinion existed, but they did not escalate to dispute situations, and we reached unanimity by means of discussion in which both viewpoints were appreciated. This compromise solution is typical of group work, and it worked well in our case" (3:101).

> "I tried to be active and express my own opinions – and I did it even sharply at times, to provoke discussion. But the deeper understanding of assessment was not reached very well

in our team. We all were perhaps lacking courage to bring our personal views into the discussions, and we preferred to go along with others' opinions of assessment. We had several assessment discussions and noticed how difficult peer feedback is to receive or give. People remained aloof, and we unfortunately started to repeat ourselves" (3: 144).

Using the concepts of Nokes-Malach et al. (2015), the first citation is an example of *observational learning,* with the idea of *complementary knowledge and error correction.* In terms of the collaboration factors of Scager et al. (2016), this refers to "mutual support" and "complementing one another." The latter excerpt reflects the notion of "*fear of evaluation*" (Nokes-Malach et al., 2015), whereby some members did not want to share their thoughts because they probably were afraid of critical comments. The researchers call it "*social loafing*" if team members just agree with others' views in order to make probable contradictions easier to handle.

DISCUSSION

Promisingly Versatile Knowledge Base of Teacher Assessment Literacy

The student teachers' reflections on the *assessment knowledge base (RQ1)* indicated that they were aware of the pedagogical significance of assessment methods and strategies, the ethically oriented responsibility to support pupils' learning processes, and pedagogical content knowledge. Less attention was paid to grading and disciplinary knowledge. Their interest in formative assessment is in line with the research evidence reported in our theoretical background (e.g., Yan and Cheng, 2015; Black and Wiliam, 2018; Cañadas, 2021). For example, Wilsey et al. (2020) found in their intervention study with experienced science teachers that "several teachers developed conceptions that were more iterative, in which frequent assessment was used to inform future instruction … The shifts in … mental models … may also suggest an approach to developing more process-oriented assessment practices in schools" (pp. 136, 154).

Awareness of versatile assessment methods is crucial for TAL because different methods improve pupils' opportunities to make their learning visible. An important enhancement in the methods repertoire was digitalization, which was a novel topic for our students. They realized how digital tools can make the assessment rationale more transparent not only for pupils and teachers (see Panero and Aldon, 2016) but also for important stakeholders, particularly guardians. The power of digital tools to motivate adolescents to view assessment favorably (perhaps as "edutainment") should not be underestimated.

Our data raised a considerable number of reflections regarding assessment ethics and responsibility, in relation to not only the knowledge base but also the conceptions. As Coombs et al. (2018) conclude, issues of fairness in assessment and the equitable treatment of students within the classroom are significant aspects for preservice teachers to reflect on in PsTE. Equitable treatment was a timely topic regarding the Finnish National Agency for Education [FNAE] (2020b) reform on assessment criteria for the final assessment of basic education. Our assessment module proved to be successful in encouraging discussions on fairness, not only as a part of the knowledge base but also regarding assessment conceptions.

Assessment Conceptions in Opening Broader Views on Teacher Assessment Literacy

Regarding *assessment conceptions (RQ2),* a similar kind of pedagogically informed orientation as in RQ1 was also evident here. Preservice teachers preferred recognising pupils' needs and promoting their ability development when assessment purposes were reflected on (Brown and Hirschfeld, 2008; Xu and He, 2019). The assessment criteria or its institutional targets did not inspire them. Perhaps surprisingly, assessment's power to enable feedback was seldom mentioned, either in relation to the knowledge base or the conceptions (Harks et al., 2014). Without feedback, pure summative or formative data do not stand on their own and do not facilitate pupils' learning.

The question of preservice teachers' trust in their assessment literacy emerged also from the analysis of conceptions, which was anticipated in our second chapter with the help of DeLuca and Klinger (2010); Levy-Vered and Alhijab (2015), and Massey et al. (2020). We did not theoretically focus on self-efficacy but were able to make observations about what was included or excluded in the reflections regarding assessment confidence. There were only a few comments in the data claiming that assessment is an easy job, but not many serious complaints about its difficulty were found. The students were fully aware of assessment's broad nature ("may influence everything") and recognized the need for constant improvement in their own professional learning.

The assessment conceptions of our preservice teachers ($N = 168$) were slightly more focused on formative assessment's role in promoting learning compared to the most recent survey data of Finnish preservice teachers ($N = 287$) reported by Kyttälä et al. (2022). Both explorations focused on similar teacher groups (subject, class, and special teachers) and introduced also a group of preservice teachers that see assessment as useless, harmful, or negative (inequal grading, stress, time-consuming). The last-mentioned group needs special attention in TE.

Student Teacher Learning as a Part of Teacher Assessment Literacy

We did not anticipate that peer group reflections would be as unanimous as they were. Xu and Brown (2016) emphasize in their TALiP-model that *teacher learning (RQ3)* in a collaborative context is crucial for assessment literacy. Perhaps the time for peer discussions was too short in our course to prompt varied or contradictory opinions of assessment. Another explanation might be the assessment criteria that were used: Modeling a learning object for a digital learning environment and justifying the choices related to it were assessed at the team level but learning diaries that included reflections on the multidimensional nature of assessment were assessed individually. In this sense, our results confirm Tinoca and Oliveira's (2012) suggestion that if online

practices are designed in accordance with the formative character of assessment, they could motivate participants' professional learning and be a valuable source of feedback for their learning processes.

The students' multiprofessional backgrounds (e.g., students of special education, history, mathematics, and counseling may have drawn lots in the same peer group) may have been challenging in terms of reaching a deeper mutual understanding if the concepts in the various fields of expertise were not yet established. On the other hand, there were good examples of progressive negotiations and complementary sharing of different viewpoints. Several studies (e.g., Charteris and Dargusch, 2018; Wilsey et al., 2020; Atjonen, 2022) suggest that assessment is not an easy topic for teachers to discuss and exchange their views and practices on with colleagues. Fear of mistakes and strong feelings of autonomy may make it difficult, and people may therefore resort to "common polite agreement."

The key perspective of our assessment module was interaction, as indicated in the title of the study module. This aspect remained slightly superficial in light of the results, although peer learning is important in the development of assessment literacy. Interaction referred not only to the interaction between pupils and teachers (involvement and engagement in **Figure 3A**) but also to the cooperation between schools and homes. Preservice teachers must become familiar with educational partnership in assessment as well. Based on our research data, school-home interaction was present in the students' reflections, although some had difficulty exemplifying this concretely.

Limitations of Our Research

Although one of our research strengths was the theoretically rigorous TALiP-model in the data analysis, we focused on only three of its components (see **Figure 1**; A, B, and E) that were not integrated tightly together in the empirical analysis. The integration was not possible because the students' reflections of the "guiding framework," "macro and micro contexts," and "assessments in practice" (i.e., three layers between components A, B, and E) were not discussed in the data. An interesting research project would be to explore students' assessment literacy from all the TALiP components by means of interviews.

Our assessment module was available for a student cohort in a TE department where voluntarily participating experienced teacher educators wanted to improve preservice teachers' assessment knowledge. Therefore, the convenience sample may question how the results can be used in other TE units. However, we were able to identify theoretical support for our efforts to strengthen student teachers' TAL (see e.g., DeLuca et al., 2013, 2019; Booth et al., 2016; Hill et al., 2017; Wyatt-Smith et al., 2017) and succeeded in describing interesting conceptions of assessment in a way that could be visible for other teacher educators as well.

Despite the presence of the TALiP-model and its conceptual illustrations in the article of Xu and Brown (2016), it sometimes required careful re-reading of the student accounts to be sure which category of knowledge base or conceptions was suitable.

On the other hand, qualitative analysis requires continued close reading in parallel with theory to fine-tune the classifications. In some cases, the students had produced longer accounts to clarify their thoughts or referred to a supplementary reply of another DA. By means of negotiations as a group of four authors of the article, we were able to become more convinced of the coherent line of interpretation.

One practical concern relates to working online due to COVID-19. Was something essential gained or missed due to distant studies as compared to "normal" lecturing, workshops, and face-to-face studies? Several surveys and practical experiences during the pandemic period indicate that some students strongly prefer personal contact in their learning. Technology-mediated peer group working may have reduced the extent and quality of the discussions in which people were not personally familiar with each other. On the other hand, the staff and students were well equipped with distance teaching and learning already in autumn 2020. We cannot indicate any significant differences in the results compared to autumn 2021, when the same assessment module was carried out with another cohort of students, and conventional classroom teaching and learning (in parallel with the remote option) was also available.

Our conclusion is that PsTE experiments need not to be resource-intensive to open preservice teachers' eyes to the various pedagogical inputs and outputs of assessment. This is in line with our previous experiences of the same TE unit (Äikäs et al., 2020; Atjonen et al., 2022) where the assessment module reported in this article was carried out. To encourage preservice teachers to enhance their assessment literacy, authentic assessment experiments with pupils must be available during teaching practice periods. The kinds of practical actions with pupils were not included in our assessment module. By means of practical experiments, more light could be shed on students' understanding of how learning objectives and assessment strategies coincide.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

Ethical approval was not provided for this study on human participants because the preservice teachers gave their written permission to use their learning diaries as research data. The data were fully anonymized and did not contain particularly sensitive data. Intervention, as such, did not require any permission, but it belonged to the autonomy of the staff members that put it into practice as a part of the research-based TE program. Based on these reasons and the principles of the UEF Ethics Committee, approval was not needed. The participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s)

for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

SP, PA, and SK contributed to the key conception, empirical content, and design of the study. SP organized the database for the analysis and wrote a part of the Results section of the manuscript. PA and SP performed the deductive and inductive analyses of data and discussed its complicated cases with SK and PR. PA wrote the first draft of the manuscript. PR elaborated the interpretations from the viewpoint of student teachers' teaching practice. All authors contributed to manuscript revision, read, and approved the submitted version.

REFERENCES

Äikäs, V., Vellonen, V., Lappalainen, K., Atjonen, P., and Holopainen, L. (2020). ""Oppilaslähtöisen arvioinnin ja tuen edistäminen erityisopettajien opetusharjoittelussa," [Promotion of pupil-centred assessment and support in the teaching practice of special education teachers]," in *Mahdoton Inkluusio? Tunnista Haasteet ja Mahdollisuudet*, eds M. Takala, A. Äikäs, and S. Lakkala (Jyväskylä: PS-kustannus), 109–139.

Atjonen, P. (2017). ""Arviointiosaamisen kehittäminen yleissivistävän koulun opettajien koulutuksessa - Opetussuunnitelmatarkastelun virittämiä näkemyksiä," [Development of teacher assessment literacy in comprehensive schools – Views from the curriculum analysis]," in *Kriteerit Puntarissa*, Vol. 74, eds V. Britschgi and J. Rautopuro (Turku: Suomen kasvatustieteellinen seura), 132–169.

Atjonen, P. (2022). ""Perusopetuksen opettajat arviointiosaamistaan kuvaamassa," [Teachers of basic education describe their assessment literacy]," in *Puheenvuoroja Arvioinnista ja Sen Kehittämisestä*, ed. P. Atjonen (Joensuu: University of Eastern Finland).

Atjonen, P., Laivamaa, H., Levonen, A., Orell, S., Saari, M., Sulonen, K., et al. (2019). *"Että Tietää Missä on Menossa": Oppimisen ja Osaamisen Arviointi Perusopetuksessa ja Lukiokoulutuksessa. ["So that We Know Where We Stand" Assessment of Learning and Competence in Basic Education and General Upper Secondary Education] Julkaisuja 7*. Helsinki: Kansallisen koulutuksen arviointikeskuksen

Atjonen, P., Ruotsalainen, P., Kontkanen, S., and Pöntinen, S. (2022). ""Opettajaksi opiskelevat formatiivista arviointia toteuttamassa," [Student teachers in putting formative assessment into practice]," in *Puheenvuoroja Arvioinnista ja Sen Kehittämisestä*, ed. P. Atjonen (Kuopio: University of Eastern Finland).

Bengtsson, M. (2016). How to plan and perform a qualitative study using content analysis. *Nurs. Open* 2, 8–14. doi: 10.1016/j.npls.2016.01.001

Black, P., and Wiliam, D. (2018). Classroom assessment and pedagogy. *Assess. Educ.* 25, 551–575.

Booth, B., Dixon, H., and Hill, M. (2016). Assessment capability for New Zealand teachers and students: challenging but possible. *Set Res. Inf. Teach.* 1, 28–35. doi: 10.18296/set.0043

Brown, G., and Gao, L. (2015). Chinese teachers' conceptions of assessment for and of learning: six competing and complementary purposes. *Cogent Educ.* 2:993836. doi: 10.1080/2331186X.2014.993836

Brown, G., and Hirschfeld, G. (2008). Students' conceptions of assessment: links to outcomes. *Assess. Educ.* 15, 3–17. doi: 10.1080/09695940701876003

Brown, G., Gebril, A., and Michaelides, M. (2019). Teachers' conceptions of assessment: a global phenomenon or a global localism. *Front. Educ.* 4:16. doi: 10.3389/feduc.2019.00016

Cañadas, L. (2021). Contribution of formative assessment for developing teaching competences in teacher education. *Eur. J. Teach.* 1–17. doi: 10.1080/02619768.2021.1950684

Cauley, K. M., and McMillan, J. H. (2010). Formative assessment techniques to support student motivation and achievement. *Clear. House* 8, 1–6. doi: 10.1080/00098650903267784

Charteris, J., and Dargusch, J. (2018). The tensions of preparing pre-service teachers to be assessment capable and profession-ready. *Asia Pac. J. Teach. Educ.* 46, 354–368. doi: 10.1080/1359866X.2018.1469114

Coombs, A., DeLuca, C., LaPointe-McEwan, D., and Chalas, A. (2018). Changing approaches to classroom assessment: an empirical study across teacher career stages. *Teach. Teach. Educ.* 71, 134–144. doi: 10.1016/j.tate.2017.12.010

DeLuca, C., and Bellara, A. (2013). The current state of assessment education: aligning policy, standards, and teacher education curriculum. *J. Teach. Educ.* 64, 356–372. doi: 10.1177/0022487113488144

DeLuca, C., and Klinger, D. A. (2010). Assessment literacy development: identifying gaps in teacher candidates' learning. *Assess. Educ.* 17, 419–438. doi: 10.1080/0969594X.2010.516643

DeLuca, C., Chavez, T., Bellara, A., and Cao, C. (2013). Pedagogies for preservice assessment education: supporting teacher candidates' assessment literacy development. *Teach. Educ.* 48, 128–142. doi: 10.1080/08878730.2012.760024

DeLuca, D., Willis, J., Cowie, B., Harrison, C., Coombs, A., Gibson, A., et al. (2019). Policies, programs, and practices: exploring the complex dynamics of assessment education in teacher education across four countries. *Front. Educ.* 4:132. doi: 10.3389/feduc.2019.00132

Deneen, C., and Brown, G. (2016). The impact of conceptions of assessment on assessment literacy in a teacher education program. *Cogent Educ.* 3:1. doi: 10.1080/2331186X.2016.1225380

Finnish National Agency for Education [FNAE] (2014). *Perusopetuksen Opetussuunnitelman Perusteet [The National Core Curriculum for Basic Education 2014]*.Helsinki: Opetushallitus. Määräykset ja ohjeet 96.

Finnish National Agency for Education [FNAE] (2020a). *Oppilaan Oppimisen ja Osaamisen Arviointi Perusopetuksessa [Assessment of Pupils' Learning and Competence in Basic Education]. Changes to the "National Core Curriculum for Basic Education 2014" FNAE10.2.2020*. Helsinki: Finnish National Agency for Education. Available online at: https://www.oph.fi/sites/default/files/documents/perusopetuksen-arviointiluku-10-2-2020_1.pdf (Accessed October 8, 2021)

Finnish National Agency for Education [FNAE] (2020b). *Perusopetuksen Päättöarvioinnin Kriteerit [Criteria for Final Assessment in Basic Education]. FNAE, 31.12.2020*. Helsinki: Finnish National Agency for Education.

Harks, B., Rakoczy, K., Hattie, J., Besser, M., and Klieme, E. (2014). The effects of feedback on achievement, interest and self-evaluation: the role of feedback's perceived usefulness. *Educ. Psychol.* 34, 269–290. doi: 10.1080/01443410.2013.785384

Hamodi, C., López-Pastor, V., and López-Pastor, A. (2016). 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). *Eur. J. Teach. Educ.* 40, 171–190. doi: 10.1080/02619768.2017.12 81909

Herppich, S., Praetorius, A.-K., Förster, N., Glogger-Frey, I., Karst, K., Leutner, D., et al. (2018). Teachers' assessment competence: integrating knowledge-, process-, and product-oriented approaches into a competence-oriented conceptual model. *Teach. Teach. Educ.* 76, 181–193. doi: 10.1016/j.tate.2017.12. 001

Hill, M., Ell, F., and Eyers, G. (2017). Assessment capability and student self-regulation: the challenge of preparing teachers. *Front. Educ.* 2:21. doi: 10.3389/feduc.2017.00021

Kyngäs, H., and Kaakinen, P. (2020). "Deductive content analysis," in *The Application of Content Analysis in Nursing Science Research*, eds H. Kyngäs, K. Mikkonen, and M. Kääriäinen (Cham: Springer), doi: 10.1007/978-3-030-30199-6_3

Kyttälä, M., Björn, P., Rantamäki, M., Lehesvuori, S., Närhi, V., Aro, M., et al. (2022). Assessment conceptions of Finnish pre-service teachers. *Eur. J. Teach. Educ.* 45. doi: 10.1080/02619768.2022.2058927 [Epub ahead of print].

Levy-Vered, A., and Alhijab, N. (2015). Modelling beginning teachers' assessment literacy: the contribution of training, self-efficacy, and conceptions of assessment. *Educ. Res. Eval.* 2, 378–406. doi: 10.1080/13803611.2015.1117980

Li, L. (2019). Using game-based training to improve students' assessment skills and intrinsic motivation in peer assessment. *Innov. Educ. Teach. Int.* 56, 423–433. doi: 10.1080/14703297.2018.1511444

Massey, K., DeLuca, C., and LaPointe-McEwan, D. (2020). Assessment literacy in college teaching: empirical evidence on the role and effectiveness of a faculty training course. *J. Educ. Dev.* 39, 209–238. doi: 10.3998/tia.17063888.0039.109

Mellati, M., and Khademi, M. (2018). Exploring teachers' assessment literacy: impact on learners' writing achievements and implications for teacher development. *Austr. J. Teach. Educ.* 43, 1–18. doi: 10.14221/ajte.2018v43n6.1

Neuendorf, K. A. (2019). "Content analysis and thematic analysis," in *Research Methods for Applied Psychologists: Design, Analysis and Reporting*, ed. P. Brough (New York, NY: Routledge), 211–223.

Nieminen, J. H., Atjonen, P., and Remesal, A. (2021). Parents' beliefs about assessment: a conceptual framework and findings from Finnish basic education. *Stud. Educ. Eval.* 71:101097. doi: 10.1016/j.stueduc.2021.101097

Nokes-Malach, T., Richey, E., and Gadfil, S. (2015). When is it better to learn together? Insights from research on collaborative learning. *Educ. Psychol. Rev.* 27, 645–656. doi: 10.1007/s10648-015-9312-8

Panero, M., and Aldon, G. (2016). How teachers evolve their formative assessment practices when digital tools are involved in the classroom. *Digit. Exp. Math. Educ.* 2, 70–86. doi: 10.1007/s40751-016-0012-x

Pastore, S., and Andrade, H. (2018). Teacher assessment literacy: a three-dimensional model. *Teach. Teach. Educ.* 84, 128–138.

Scager, K., Boonstra, J., Peeters, T., Vulperhorst, J., and Wiegant, F. (2016). Collaborative learning in higher education. Evoking positive interdependence. *CBE Life Sci. Educ.* 15:ar69. doi: 10.1187/cbe.16-07-0219

Schneider, C., and Bodensohn, R. (2017). Student teachers' appraisal of the importance of assessment in teacher education and self-reports on the development of assessment competence. *Assess. Educ.* 24, 127–146. doi: 10.1080/0969594X.2017.1293002

Stiggins, R. (1991). Assessment literacy. *Phi Delta Kappan* 72, 534–539.

Stiggins, R. (2014). Improve assessment literacy outside of schools too. *Phi Delta Kappan* 96, 65–72. doi: 10.1177/0031721714553413

Timperley, H. (2009). Using assessment data for improving teaching practice. *Austr. Coll. Educ.* 8, 21–27.

Tinoca, L., and Oliveira, I. (2012). Formative assessment of teachers in the context of an online learning environment. *Teach. Teach.* 19, 214–227. doi: 10.1080/13540602.2013.741836

Volante, L., and Fazio, X. (2007). Exploring teacher candidates' assessment literacy: implications for teacher education reform and professional development. *Can. J. Educ.* 30, 749–770. doi: 10.2307/20466661

Willis, J., Adie, L., and Klenowski, V. (2013). Conceptualising teachers' assessment literacies in an era of curriculum and assessment reform. *Austr. Educ. Res.* 40, 241–256. doi: 10.1007/s13384-013-0089-9

Wilsey, M., Kloser, M., Borko, H., and Rafanelli, S. (2020). Middle school science teachers' conceptions of assessment practice throughout a year-long professional development experience. *Educ. Assess.* 25, 136–158. doi: 10.1080/10627197.2020.1756255

Wyatt-Smith, C., Alexander, C., Fishburn, D., and McMahon, P. (2017). Standards of practice to standards of evidence: developing assessment capable teachers. *Assess. Educ.* 24, 250–270. doi: 10.1080/0969594X.2016.122 8603

Xu, Y., and Brown, G. (2016). Teacher assessment literacy in practice: a reconceptualization. *Teach. Teach. Educ.* 58, 149–162. doi: 10.1016/j.tate.2016. 05.010

Xu, Y., and He, L. (2019). How pre-service teachers' conceptions of assessment change over practicum: implications for teacher assessment literacy. *Front. Educ.* 4:145. doi: 10.3389/feduc.2019.00145

Yan, Z., and Cheng, E. (2015). Primary teachers' attitudes, intentions and practices regarding formative assessment. *Teach. Teach. Educ.* 45, 128–136. doi: 10.1016/j.tate.2014.10.002