Author(s): Ketonen, Laura; Körkkö, Minna; Pöysä, Sanni

Title: Authentic assessment as a support for student teachers' reflection

Year: 2023

Version: Published version

Please cite the original version:

Ketonen, L., Körkkö, M., & Pöysä, S. (2023). Authentic assessment as a support for student
teachers' reflection. European Journal of Teacher Education, Early online.
https://doi.org/10.1080/02619768.2023.2229004

Authentic assessment as a support for student teachers' reflection

Laura Ketonen, Minna Körkkö & Sanni Pöysä

Published online: 28 Jun 2023.

Submit your article to this journal ⬚

Article views: 30

View related articles ⬚

View Crossmark data ⬚

Routledge
Taylor & Francis Group

Authentic assessment as a support for student teachers' reflection

Laura Ketonen [a,b], Minna Körkkö [c] and Sanni Pöysä [a]

aDepartment of Teacher Education, University of Jyväskylä, Jyväskylä, Finland; bFinnish Institute for Educational Research, University of Jyväskylä, Jyväskylä, Finland; cFaculty of Education, University of Oulu, Oulu, Finland

ABSTRACT

Assessment and feedback guide learning. In light of this, the assessment practices of teacher education have received little research. This qualitative study examines student teachers' discussions in a study unit built on authentic assessment practices: self- and peer assessment of videotapes from authentic performance with research-based rubrics. The aim is to investigate whether authentic assessment supports student teachers' reflection and, if so, how. The findings show that authentic assessment led students frequently to reflection, and in most cases reflective discussions were induced by students' self-criticism. We deduced that the encouraging feedback culture between students and between students and the teacher enabled students to be open about their self-critical observations. According to the findings, building a study unit on authentic assessment is a promising way to guide students to reflect on theory and practice and to learn skills that are essential for their future profession.

Introduction

Assessment is a key component of learning (Rodríguez-Gómez and María Soledad 2015). Learning is powerfully guided by the combination of summative and formative assessment practices, and therefore the alignment of assessment with learning objectives and instruction is vital (Biggs and So-Kum Tang 2011; Black and Wiliam 2009). When the instruction or objectives of learning change, the assessment needs to be re-examined as well. Teacher education is currently being transformed towards more research- and practiced-based procedures (Afdal and Spernes 2018; Matsumoto-Royo and Soledad Ramírez-Montoya 2021), but despite this, the discussion about the assessment processes in teacher education is scarce. Studies reporting renewed practices of teacher education seldom consider, or even mention, the aspect of assessment (Matsumoto-Royo and Soledad Ramírez-Montoya 2021). Therefore, there is a need for more thorough

descriptions and analyses of assessment processes in teacher education (Matsumoto-Royo and Soledad Ramírez-Montoya 2021).

The research on assessment and feedback has taken major strides in the past decade; the focus has shifted from teachers towards students' active participation (e.g. Carless and Boud 2018; Nieminen and Tuohilampi 2020; Winstone et al. 2017). The rationale behind the change of paradigm is that even though teachers should provide the most fruitful feedback, it does not help students if they do not engage with it and if they are not prepared to use it for their development. The intent of higher education is to prepare students for a professional world, one in which they need to evaluate their own performance and seek feedback. It has been suggested that even if the prevalent feedback practices of the future profession are harmful, it would be useful to let students become accustomed to those practices during their studies and learn how to react to them (Dawson, Carless, and Pui Wah Lee 2021). Higher education should not make students dependent on their teachers' feedback, but instead, during their studies, it should let them develop their feedback literacy so that they can act independently by the time they start in their first workplace. This concerns teacher education even more than other professional education, because future teachers are not only required to evaluate their own performance, but also their students' performance. Hence, we need research about those practices that help student teachers develop their ability to seek, evaluate and use feedback. Authentic assessment could provide opportunities for this.

Authentic assessment signifies using assessment tasks, standards, and feedback practices that have been influenced from work life and, thus, have authentic elements, for example, feedback concerning work-related skills or feedback coming in a form that is typical for the profession. According to research, authentic assessment has a positive impact on student learning, commitment and motivation for learning and metacognition (Villarroel et al. 2018). The intent of authentic assessment is to bind university learning and work-related skills together. In teacher education, a natural context for authentic assessment is guided practical training, which is highly authentic in itself. Practical training provides students with an opportunity to observe and evaluate their own and their peers' performance. In the present study, we explore the potential of authentic assessment for student teachers' learning.

Much teacher learning takes place through reflection (Korthagen and Vasalos 2005; Körkkö, Kyrö-Ämmälä, and Turunen 2016). As a distinction from most professional education, in teacher education, students come to teacher education with years of school experience in a pupil's role and with conceptions of teaching and learning (Körkkö, Kyrö-Ämmälä, and Turunen 2016). Therefore, teacher education is not only about learning, but also about transforming pre-existing incorrect or outdated conceptions. In addition, pure memorising or understanding of educational concepts and theories does not guarantee the ability to apply those to practice, but it is through reflection that theories are attached to practice.

Intuitively, one could expect that authentic assessment would encourage student teacher reflection, but does this actually work? In the present study, we explore this by examining student teachers' discussions in a learning module built on authentic assessment practices. The specific aims of the study are to understand the following:

(1) Does authentic assessment setting guide students to reflection?
(2) What elements of the authentic assessment setting lead students to reflection?

Authentic assessment for professional development

Authenticity is considered one of the most fruitful characteristics of assessment design (Villarroel et al. 2018). Authentic assessment can be understood as having three dimensions: realism, cognitive challenge, and evaluative judgement. *Realism* of assessment is sought by using tasks, problems, and assessing competences that are similar to those of work life and relevant outside the university. The organisation of assessment itself can entail authentic elements, for example, collaboration or peer feedback (Dawson, Carless, and Pui Wah Lee 2021; Villarroel et al. 2018). *Cognitive challenge* relates to performing tasks that require higher-order thinking, problem-solving, and decision making. In work life, these tasks tend to be complex, requiring not only remembering and understanding the information, but also applying it in a context-dependent way to the practice. Authentic assessment aims to provide students with similar challenges. *Students' evaluative judgement* can be pursued by having feedback dialogues, discussing exemplars, using formative self- and peer assessment, and basing these on transparent criteria (Tai et al. 2018). Thus, students can learn to evaluate their own performance and constitute conceptions of quality. These three dimensions will be further explained with examples in the method section, where we introduce the organisation of the study unit. The possibilities to implement authentic assessment in teacher education have been discussed earlier (Darling-Hammond and Snyder 2000), but the focus was on the summative side, which was typical for the assessment research of the time. There is a need to understand more deeply the formative side, that is, the mechanisms through which the authenticity of assessment influences student teachers' learning.

Reflection in teacher education

Previous studies consider reflection to be a key element of student teachers' professional development (Korthagen and Vasalos 2005; Körkkö, Kyrö-Ämmälä, and Turunen 2016). Therefore, many teacher education programs follow a reflective approach, in which the aim is to educate teachers who can integrate theoretical knowledge into their practice and critically examine their experiences and actions (Dewey 1933; Schön 1983). Through a reflective process, students develop their practical knowledge of teaching, which is then used in teacher's actual work (Levin and He 2008).

In the past few decades, the concept of reflection has been defined multiple ways and from different perspectives; therefore, there is no one way to describe reflection. Many scholars base their thinking on the work of Dewey (1933), who defined reflection as a systematic way of thinking about practice to improve it. Since Dewey's pioneer work, the definitions of reflection have been revised and broadened; however, different definitions usually share the same basic principles: reflection is firmly situated in practice, and it evolves through a cyclical and progressive process where a teacher looks back on action, conceptualises it, seeks multiple perspectives, and plans new action (e.g. Kolb 1984; Mezirow 1991; Schön 1983). Moreover, as Dewey highlighted, the social aspect of

reflection is considered essential because interaction with others significantly promotes reflection (Christ, Arya, and Ming Chiu 2014; Korthagen and Vasalos 2005).

The theoretical literature includes different hierarchical qualities of reflection. The most common way is to divide reflection into those levels of thinking that represent either more superficial or more advanced thinking (e.g. Jay and Johnson 2002; Van Manen 1977). The lowest level of thinking can be called descriptive or technical, and it includes merely mentioning the problem. When thinking evolves from a superficial level, it includes consideration of alternative views and questioning one's assumptions, thus entailing a more critical stance.

In teacher education, the aim is to promote student teachers' critical reflection skills, that is to help them to look at their teaching from alternative viewpoints, question their way of thinking and connect their actions to educational theory (Toom et al. 2010). Previous studies indicate that questioning, criticising, and especially changing one's firmly held beliefs and actions is challenging for student teachers, which highlights the role of guidance in promoting development of reflection skills (Körkkö 2021; McGarr and McCormack 2014). Student teacher reflection can be supported and guided through different artefacts such as portfolio writing and video recording. While portfolios might encourage more superficial reflection (Chye et al. 2019), previous studies have shown the effectivity of video in fostering adoption of a critical stance and therefore, more and more studies focus on using video as a tool for teacher learning (e.g. Danielowich 2014; Körkkö 2020; Stockero, Rupnow, and Pascoe 2017).

Guided teaching practicum periods seem to be essential arenas where collaborative discussions with supervisors and peers promote student teachers' reflection and construction of their teacher identities (Körkkö, Kyrö-Ämmälä, and Turunen 2016). Findings of Paksuniemi et al. (2021) show the essential role of the supervisor in creation of a safe learning environment where students are encouraged to engage in dialogue with each other. Supervisors can promote student teachers' learning and professional development by focusing on issues behind the students' actual and visible teaching behaviours, such as their theoretical knowledge, beliefs, motivation and emotions (Korthagen 2017; Körkkö 2021). This helps student teachers to become more aware of multiple reasons of their actions, which further enables adoption of a critical stance to teaching.

Mezirow's approach to reflection

In the present study, we base the observation of the quality of reflection on Mezirow's work. There is a consensus about the importance of reflectivity and reflective thinking in teachers' professional learning, but instruments to measure the depth of that thinking are scarce (Beauchamp 2015). Many reflection models developed for teacher education (e.g. Jay and Johnson 2002; Korthagen and Vasalos 2005; Valli 1997) do not explicitly define the difference between non-reflection and reflection, and therefore they do not serve our study's needs. Mezirow's transformative learning theory, however, has been used to develop instruments that determine the distinct levels of reflection (Henriette and Poell 2016).

Mezirow's conceptualisation of reflection evolved over the years. In the current paper, we built on the reflection continuum developed by Peltier et al. (2005) (see also Kember et al. 2000), which is based on Mezirow's paper from 1991 (Henriette and Poell 2016). Peltier's model describes four levels of reflectivity: habitual action, understanding,

Table 1. The nonreflection – reflection continuum by Peltier et al. (2005).

Nonreflection/Surface Learning <—————-> Reflection/Deep Learning			
Habitual Action	Understanding	Reflection	Intensive Reflection
Minimal thought and engagement, correlated with a surface approach to learning – specific tasks are treated as unrelated activities, memorisation is emphasised and which embodies an attitudinal state of unreflectiveness.	Focuses on comprehension without relation to one's personal experience or other learning situations. Book learning is understanding-oriented in that the learner needs only comprehend the read materials. Most of what is learned stays within the boundaries of pre-existing perspectives.	Learning is related to personal experience and other knowledge. Reflection also involves challenging assumptions, seeking alternatives, and identifying areas for improvement. Shows active and conscious engagement, characteristics commonly associated with a deep approach to the learning.	Intensive reflection is at the highest level of the reflective learning hierarch, and learners become aware of why they think, perceive, or act as they do. Learners might alter or even completely change firmly held beliefs and ways of thinking. Intensive reflection is thus seen as involving a change in personal beliefs.

reflection, and intensive reflection. These are explained in Table 1. We chose this model because of its explicit conceptualisation of reflection and the clear distinction between reflection and non-reflection. Also, the model is inclusive of various aspects of reflection: cognitive, social, and emotional. Our study corroborates previous studies based on Mezirow's ideas. The scales inspired by the work of Mezirow have mostly been used with written artefacts, interviews, and group interviews (Henriette and Poell 2016), but in the current research, we apply the scale with authentic learning discussions.

Method

The organisation of the study unit using authentic assessment

The current study was conducted as part of teacher education in a Finnish university. Finnish teacher education follows a research-based approach, which means that teaching is based on the latest research and student teachers learn academic and research skills, such as analytical and critical thinking, through different activities and assignments; moreover, student teachers practice research and write bachelor's and master's theses (Lauriala 2013; Toom et al. 2010). In Finland, all teachers graduate with a master's degree.

The study unit of the present study was the product of a development project that was funded by Finland's Ministry of Education and Culture for the years 2018–2021. The study unit was based on the VOPA procedure (Pöysä et al. 2021), which focuses on the role of classroom interaction as a basis for teaching and assessment (the Finnish name behind the abbreviation VOPA can be translated as 'classroom interaction as a basis for teaching and assessment'). The VOPA procedure focuses on four themes: 1) teacher sensitivity and positive climate, 2) classroom organisation and motivating students, 3) dialogicity, and 4) feedback. Themes are separate in such a way that it is possible to concentrate on only one theme or to choose more than one theme. In the present study unit, three out of four units were targeted: 1) classroom organisation and motivating students, 2) dialogicity, and 3) feedback (Figure 1).

The structure of the VOPA procedure has a basis in research (Pöysä et al. 2021). Its key element was research-based rubrics composed for each theme. The rubrics present the

central components of the interaction that are theoretically relevant to teaching and assessment. In addition, the rubric provides phased descriptions on how the participants can improve their teaching and assessment skills with respect to specific components from the beginner state (phase 1) up to the competent state (phase 3; the Finnish rubrics can be found in Pöysä, Pakarinen, Ketonen et al. 2021). Based on the structure of the VOPA procedure, each theme included three steps (Figure 1). The first step is *Grounding*, that is, a 90-minute meeting in which the theoretical background of the theme is studied, and the usage of rubric is practiced via discussions and video examples. The second step is *Execution*, that is, the time when the participants record one lesson of their own teaching and watch that in 15- to 20-minute sequents. While watching, the participants compare their own teaching with the contents of the rubric and choose a sample (approximately five minutes long) of a good practice with respect to some aspect of the rubric. The third step is *Joint discussion*, that is, a 90-minute meeting in which the participants share their sample video and watch other participants' videos. Samples are discussed, and the participants are encouraged to share and receive feedback based on the rubric of the theme. The features of authentic assessment of the study unit are listed in Table 2.

Participants and data gathering

The participants in the present study were 15 student teachers and their university teacher. The group of student teachers was heterogeneous regarding their educational background (major subject either educational sciences or special education), their age (mean 22 years, range 20–27 years), and their number of passed course credits in the university (number of credits at the beginning of academic year, mean 155 credits, $SD =$ 35 credits, range = 80–200 credits). The participants were selected based on a study module (25 credits) that they had chosen for the year that the data were collected. Information concerning the study was given to the student teachers via discussions and in written format. They were informed about their rights and explained that not joining the study would not influence their studies. All 15 student teachers participating in this specific study module were asked to participate, and everyone gave their written consent voluntarily. The teacher participating in the present study was responsible for teaching most of the contents for the study module. The educational background of the teacher was a Master of Educational Sciences, along with a minor in special education. She had few years of experience as a university teacher, and she knew the VOPA procedure well.

The study unit from which the data were collected was formed of six 90-minute meetings scheduled over 20 weeks (Figure 1). However, the data used for the present study were collected only from the 'joint discussion' meetings, in which the student teachers shared their own video examples and had shared discussions based on those.
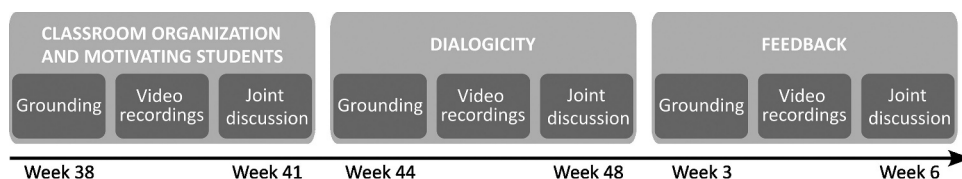

Figure 1. The organisation of the study unit.

Table 2. Elements of authentic assessment in the study unit.

Dimension	Subfeature	Elements in the study unit
Realism	Real context	o Students show authentic videos from their teaching practice. o The training model reminds that of professionals: it is also used with in-service teachers.
	Task similar to that of real and/or professional life	o Student teachers' real performances as teacher are assessed. o The focus is on interaction, which is a teacher's central professional skill. o The skills needed in and practiced with assessment tasks—noticing and interpreting—are needed in the teacher's job.
Cognitive challenge	High-order thinking skills	o The task concerns classroom interaction that is complex. o The task needs applying theory to practical situations and analysing and assessing one's own and others' performance.
Evaluative judgement	Understanding criteria and standards for good performance	o Research-based rubrics illustrate what good quality interaction consist of. o Video clips introduce moments of students' best performance. They demonstrate what good interaction might look like.
	Learning to judge their own performance	o Students evaluate their performance with the rubric as they watch their own video before the session and choose a moment that illustrates elements of good interaction. o Teacher's and peers' feedback assists students in calibrating their own views with others'.
	Formative assessment	o Students receive teacher's and peers' feedback in learning conversations.

The meetings were video recorded and transcribed, and the transcriptions formed the data for the present study.

Analysis

The analysis was qualitative and theory driven. As a first step of analysis, the first researcher watched the videotapes and transcribed the teacher's and student teachers' discussions. The length of the transcribed discussions that concerned the videotapes – practical issues excluded – was 18 300 words in Finnish. As a second step of the analysis, the transcriptions were divided into analysis units. The unit of analysis was the participant's full speaking turn, even when it contained more than one aspect. Only if there was a lengthy silence after the participants turn ended and then they continued with another topic were the units considered separate. The learning conversations were mostly organised, and the speakers used long turns to explain their thoughts carefully. Short comments that supported or supplemented the speaker's story were not counted as speaking turns if they did not change the course of discussion.

Coding was the third step in the analysis; this was based on the levels of reflection (Peltier et al. 2005) that are introduced in Table 1. The first researcher, who knew the data and Peltier's framework thoroughly, labelled each unit of analysis by using the preliminary code list that she had derived from theory, creating new codes if the preliminary ones did not apply. The preliminary and final codes are introduced in Table 3. The coding was an iterative process. After the researcher had coded the whole data, she examined the codes and combined and renamed them. Afterward, she coded the data again with renewed

Table 3. Preliminary and final codes.

Level of reflection	Preliminary codes	Final codes	Explanation of the final code
Habitual action		-	
Understanding	Comprehension of theory (or material) Book learningNoticing (elements of theory, or connecting the theory and situation)Book learningNoticing (elements of theory, or connecting the theory and situation) Book learning Noticing (elements of theory, or connecting the theory and situation)	Beginning	Commentor question that begins the conversation after the video clip.
		Question	Question about the theme or relating phenomena.
		Teacher's guiding question	Teacher's leading question that intends to guide the discussion or observations towards a certain direction, for example, return it to the theory.
		Question about background information	Question that seeks for additional information about the situation or student group on the video clip before or after the actual clip.
		Background information	Information about the situation or student group that does not necessarily show up on the video.
		Noticing	Noticing elements of interaction on the video and connecting theory and observation.
		Assessing	Evaluative commentthat does not directly relate to the topic. For example, feedback about student teachers' general skills.
		Considering the phenomena	Considering the phenomena and theory on the general level, not only in the video clip's situation.
Reflection	Personal experiencePrior knowledgeChallenging assumptionsSeeking alternativesIdentifying areas for improvementPrior knowledgeChallenging assumptionsSeeking alternativesIdentifying areas for improvement Prior knowledge Challenging assumptions Seeking alternatives Identifying areas for improvement	Alternatives	Considering alternative ways of thinking, acting, or interpreting the situation.
		Teacher confirms alternatives	Teacher shows that she agrees with the alternative way of thinking.
		Areas for improvement	Considerations about how the practice or way of acting would be developed (only noticing the need of improvement is not reflection).
		Personal experience	Student expresses their own personal experience about the situation (only describing the situation is not reflection).
		Challenging assumptions	Commentthat challenges assumptions (that the commentor considers to exist).
Intensive reflection	Why I think the way I doChange in prior beliefs or ways of thinkingChange in prior beliefs or ways of thinking Change in prior beliefs or ways of thinking	-	
Other		Other	Commentoutside the topic.

codes, creating new ones, if appropriate. After three rounds of coding and adjusting the codes, further rounds did not cause any changes. In this phase, the first researcher conducted peer debriefing (Onwuegbuzie and Leech 2007) with the second researcher by introducing the codes and findings, here with a particular focus on extracts that she

had coded as reflection. Peer debriefing was used instead of peer coding because we considered the co-development of coding scheme essential alongside with testing it. Theories of reflection are the second researcher's expertise, and consultation on the coding choices, such as classification of 'noticing' as understanding-level operation, were valuable in creating understanding reflection in video-based learning discussion. The second researcher agreed that the exact discussions that the first researcher had considered containing reflection indeed contained it, but she proposed changes to coding on the discussion turn level. These propositions were discussed until agreement between the researchers has been found.

After coding each turn, the first researcher formed graphs illustrating the course of discussions regarding the level of reflection. With the graphs and transcripts, she examined the turns that preceded the moments of reflection to find what induced them. Afterward, she examined what came after the episodes that started from incentive and led to reflection. The sequence following a single case does not explain its emergence, but it mirrors the culture of the group and assists in understanding why similar episodes emerge – or do not emerge – in the future. As a final step, the researcher used member checking by discussing the findings with the teacher of the study (Onwuegbuzie and Leech 2007).

Results

The first aim of the study was to examine whether the learning discussions of the study unit built on authentic assessment contained reflection. With identification of reflection, we used the reflection – nonreflection continuum by Peltier, Hay, and Drago (2005). During three 90-minute sessions, the student teachers presented 13 videotapes and discussed each of them after seeing the video clip. At a minimum, all the discussions were on the understanding level. Most student teachers' and the teacher's discussion turns were on the level of understanding, especially on noticing. Another typical discussion turn on the understanding level concerned background information about the situation or the student group.

In 6 of the 13 discussions about videos, student teachers' discussions reached a level of reflection, and this was true in three discussions about two issues. The student teachers considered areas for improvement, challenged assumptions, and sought alternative ways of thinking, acting, and interpreting situations. The discussions did not reach the level of intensive reflection, which is not surprising because intensive reflection, that is, transforming one's underlying assumptions, typically takes a long time. It is possible that the student teachers changed their prior beliefs during the entire study unit, but this was not tracked in the discussions.

The second aim of the study was to identify the elements that lead student teachers to reflection. When examining what preceded and induced the reflection, we noticed that in all but one case, the reflection originated from criticism. There were nine topics of reflection, of which all but one originated from critical observation, and correspondingly, only once did a student teacher express criticism making it so that the discussion did not lead to reflection. Most often, the student teacher who showed their video was the one who expressed a critical viewpoint about it. Once, though, the critical viewpoint came from a peer. We also looked into what came after the criticism. Because criticism was

expressed regularly during the discussions, we deduced that the student teachers felt comfortable expressing it. Therefore, we examined student teachers' and the teacher educator's responses to that criticism with an intention to learn.

Other student teachers were encouraging each time criticism was expressed. They considered those aspects that supported the choices that the student teacher had made, emphasised the good they saw, and gave their perspective by sharing their 'much worse' experiences. The teacher educator let the student teachers do most of the talking, but in each case of criticism, she did one of two things: she either (1) brought up good elements of the student teacher's behaviour on the video or (2) expressed that she appreciated the student teacher's critical observation. Hence, her feedback was always encouraging, even when she agreed with the criticism. Additionally, she supported the safe environment and reflection by allowing the student teachers to express opposite views. We will demonstrate these findings with two examples: one about the student teacher's self-criticism and another where critical consideration is given by a peer. In both cases, we will present a graph about the whole discussion with extracts of the transcript. In the graphs, speaking turns about the topic are marked with 'o', questions with '?', and comments unrelated to the topic with 'x'. The teacher educator's turns are bolded. The numbers above the graph refer to the discussion turns in the transcript below. The graphs of all discussions are presented in an Appendix.

Student's self-criticism reveals the encouraging feedback culture

Most often, criticism was expressed by the student teacher showing the video. An example of such a case is shown in Figure 2 and discussed below.
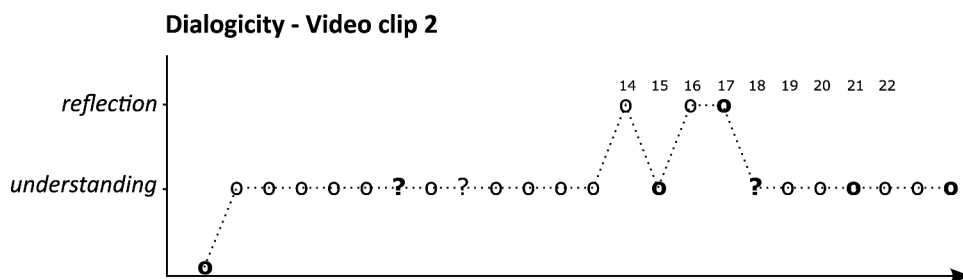
Figure 2. Graph of discussions about video clip 2 from the theme of dialogicity.

(14) Student Teacher 1: How many times can you threaten [a student with changing their seat] without doing it? I would have stopped talking about changing the seat and only had that conversation [as shown on the video]. That bothered me. I don't mean that it [conversation] would necessarily have worked, but [...] when we talked about that [changing the seat], it kept bothering me that we didn't do it.

(15) Teacher Educator: It's about what *you* notice and what *we* notice [...] When I saw the video, it wasn't my first thought that I wished that you had done

something differently. The first was that it works well, when you calmly discuss [. . .]

(16) Student Teacher 1: And then, when we filled in that student form with ST3 after this, we talked about how you don't realise how well those students work and concentrate and all that. You should boost them and give [them] positive feedback about working. You easily give it only after a challenging situation. But [. . .] you forget [the feedback], as most of them are such angels.

(17) Teacher Educator: What an insight there for you! You realised that, 'Wait a minute, I could have given more [feedback]'. Such concrete attention boosts your growth as a teacher. Therefore making such findings is excellent!

(18) Teacher Educator: Ok, how do you consider that the dialogicity succeeded in that clip?

(19) Student Teacher 1: Well, when I looked at the video, I thought the questions were maybe too direct and then, I don't know, I find weaknesseseverywhere.

(20) Student Teacher 2: You were critical in showing your video. Notice the good!

(21) Teacher Educator: Exactly! Therefore, I dared to ask you questions, since you have nothing to be embarrassed about. You can sit there being proud of yourself and listen when others praise you. This was a great clip.

The discussion concerns a video clip in which Student Teacher 1 is the teacher and holds a discussion with a young student. Before the discussion on the video clip, the student kept peeking at the results from his classmate instead of solving them himself. Student Teacher 1 had instructed him several times to concentrate on his own work. She had told him to move his desk further from his classmate, and she had threatened that if the student did not stop peeking, he would be moved to another seat.

On the video clip, Student Teacher 1 holds a constructive dialogue with the student about the point of making exercises and the disadvantages of peaking. The situation is exemplary, and she gets plenty of positive feedback about it. However, she regrets threatening the student multiple times with the seat change and not carrying out the threat. This is where the extract begins. Student Teacher 1 expresses this as an area for improvement and continues with reflecting on alternative ways of acting, particularly by not threatening the student in the first place but instead having only the constructive discussion.

The teacher educator reacts to Student Teacher 1's self-criticism by sharing her positive interpretation about the dialogue on the video clip. She emphasises the good in the situation. Student Teacher 1 interrupts her and continues by bringing up another consideration about the need for development. She had noticed that she tends to give positive feedback after challenging situations but forgets to give it when students work appropriately. Now, the teacher educator reacts by giving positive feedback about student's noticing. She compliments Student Teacher 1 for the good insight and says that such a finding allows her to develop as a teacher. Whether the teacher agrees or
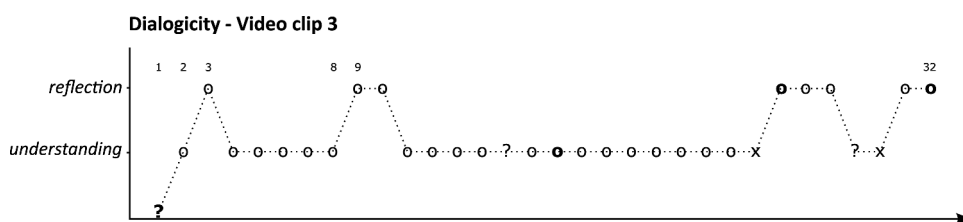
Figure 3. Graph of discussions about video clip 3 from the theme of dialogicity.

disagrees with the student teacher's self-criticism, her intervention is supportive. She either gives positive feedback about the student teacher's behaviour on the video or about her critical finding.

Other student teachers' reactions to self-criticism were also supportive. For example, Student Teacher 2 tells Student Teacher 1 to 'notice the good', and Student Teacher 3 explains how the situation was spontaneous, so Student Teacher 1 should be less critical towards her open questions because she had not planned those.

Peer's criticism leads to a debate

Only once did another student teacher suggest a topic of development. This conversation is shown in Figure 3.

In this conversation, Student Teacher 4, after giving positive feedback about several aspects of other student teacher's video clip, adds that one dimension of the rubric, encouraging climate, could have been improved. The commentleads to a lengthy debate between student teachers about how encouragement should appear in teacher's interaction; what is adequate and what is excessive. The teacher educator allows the student teachers to discuss the topic for a long time, and they reach the level of reflection several times without her direct input. However, at the very end of the conversation she gives her opinion:

(32) Teacher Educator: Ok, I want to [. . .] First, like I said, excellent reflection from you all on this video. I agree that I would not say that the climate was not encouraging or supportive. I notice that the climate was warm and supportive of participation. If we consider this from the perspective of indicators and try to identify where the support and encouragement was exactly, it is not obvious. But it does not make this clip less valuable if it did not catch [these]. Do you understand what I'm saying? A great clip. This is an especially good example of dialogicity, considering that it was only a few minutes. Very well chosen so that many elements of dialogicity are shown. But I understand the commentof Student Teacher 4, because encouragement was not shown as strongly as it is presented in the criteria [. . .]. From the climate perspective, it was highly encouraging, so the children had the courage to participate.

Table 4. Elements of authentic assessment that supported reflection.

Dimension	Subfeature	Element supporting reflection	Support for reflection
Realism & Cognitive challenge	Authentic context that requires high-order thinking skills	Videotapes show the complexity of real teaching situations	Lays basis for reflection by letting various notices to rise and several reasonable interpretations to coexist.
Evaluative judgement	Formative assessment	Student teachers give themselves (and once to their peer) critical feedback about the *task*	Leads to reflection
		Teacher and peers give encouraging feedback about the *task*	Builds encouraging feedback culture
			Encourages student teachers to share critical self-feedback
		Teacher gives positive feedback about noticing	Builds encouraging feedback culture
			Encourages student teachers to share critical self-feedback
	Modelling evaluative judgement	Teacher considers and validates different views	Models reflection
			Encourages student teachers to present divergent views.

As a closure to the whole discussion, the Teacher explains that she finds the climate on the video encouraging, but she understands Student Teacher 4's point of view because the actual indicators of an encouraging climate are implicit in the video. By saying this, the teacher validates Student Teacher 4's initiative but communicates that encouragement goes beyond explicit indicators, hence also validating the other student teachers' ideas.

Results from the perspective of authentic assessment

The purpose of this section is to draw together the main findings and to present them in the frame of authentic assessment.

Reflection originated from critical comments that were followed by opposing views, and this was supported by several elements (Table 4). The real context of teacher's work is complex, and videotapes revealed enough of that reality, laying basis for serious evaluative judgement. Moreover, the complexity of the task allowed coexistence of diverse reasonable views. The teacher also modelled reflectivity by considering student teachers' divergent views and when possible, validated them all.

Student teachers were open about their self-critical observations. This was supported by encouraging feedback culture, which comprised teacher's and peers' encouraging feedback about the task, and teacher's encouraging feedback about noticing one's needs for development.

Discussion

The aim of the present study was to examine whether and how authentic assessment supported student teachers' reflection. The student teachers showed video clips of good practices of their own authentic teaching practice, discussing the clips afterwards with rubrics that contained research-based criteria for teacher's interaction. Almost half of the discussions reached the level of reflection, which we consider a high amount because reflection was not framed as the aim of the learning discussions. The elements of authentic assessment laid the basis for reflection: the rubric made theory concrete,

provided common concepts for discussion, and described what a high-quality interaction would be like. Watching one's own video, assessing one's own performance, and choosing a video clip drove student teachers to evaluate their own performance, while watching others' video clips about good practices demonstrated what a quality might look like. The provision of peer feedback made student teachers evaluate peers' performance, and discussions about the clips led student teachers to consider and weigh different views and interpretations about the actual clips and beyond them. The discussions were never superficial – that is, at a minimum, they were on the level of understanding – and the student teachers appeared engaged, even though their performance did not influence their grades. We argue that consideration of each other's authentic professional performances was so captivating that the deliverance of grades would not have added student teachers' motivation. Rather, grades might have decreased it by lowering the authenticity of the task. This argument is supported by our experience of successfully employing the same training model with no external rewards with in-service teachers, who perceived it as fruitful for their professional development (Ketonen et al. 2021).

Critical feedback has been suggested as efficient for development (Ketonen et al. 2020), and the findings of the current study support this because in all but one case, reflection originated from criticism. Critical feedback is delicate. It can threaten one's identity and provoke defensive responses (Carless and Boud 2018) – possibly even more so when the feedback concerns one's authentic performance. For a teacher, the distribution of critical feedback while maintaining a comfortable environment is a challenge; however, both appear important. Reflection demands a safe environment because it is about expressing doubt, sharing thoughts and feelings, and discussing uncertainties (Peltier et al. 2005). The findings of the current study propose one solution to the challenge: critical feedback does not necessarily require critical commenting from others because it can be attained by using a task that encourages the participants themselves to identify areas for improvement and creating a safe climate that allows them to express critical thoughts about their own performance. An advantage of self-criticism as a form of feedback is that it emerges as already received. What is left for feedback dialogue is the joint evaluation of feedback and examination of ways to improve. The alignment of this strategy with the authentic feedback practices of the teacher's profession increases its value in teacher education. For a teacher, self-assessment is the main source of feedback, and therefore it is rational to practise such a strategy during teacher education. We argue that authentic assessment should be exploited in teacher education to evoke student teachers' internal and communal feedback processes. We acknowledge this approach's time-consuming nature, but we consider it worthwhile, as it includes several important aspects: practising to notice, practicing to reflect, practising to produce internal feedback and practising to provide collegial feedback. As the organisation of authentic assessment as described in this study requires acting in the teacher's role, we suggest it is used alongside teaching practice to support and reflect teacher students' experiences.

The role of the teacher educator in building a safe environment, facilitating dialogue, and guiding students to reflect is fundamental (Körkkö 2021; Paksuniemi et al. 2021; Peltier et al. 2005). The teacher educator in the current study consistently promoted a safe climate and reflection, although she let students do most of the reflection. Firstly, she delivered only encouraging feedback. If she agreed with student teachers' self-criticism,

she congratulated them about the important finding, and when she did not consider the criticism urgent, she brought up alternative perspectives about their performance, emphasising the good in the video. Secondly, when student teachers had divergent unresolved views, she validated both views by recognising the evidence that supported them and continued by giving and justifying her own opinion about the matter. We consider the validation of divergent views as significant. Because reflection is about challenging beliefs and testing alternative ways of thinking, it is crucial that the teacher does not imply that the aim is to find the 'right answer' but instead allows the emergence of diverse, even opposite, views. In this way, the teacher encourages student teachers to express doubt and disagreement – elements that Peltier et al. (2005) considered important for reflection. What we also considered as supporting the encouraging climate was instructing the student teachers to select video clips of good practices for joint discussion and, when providing peer feedback, guiding them to emphasise the good they saw on the video. We consider these ways of promoting a safe climate and reflection as the teacher educator's tactical choices rather than her characteristics, and we encourage other teacher educators to apply them, especially when discussing complex or delicate phenomena, such as classroom interaction.

The current study has demonstrated that authentic assessment with videotapes from teaching practice can encourage student teachers' reflections. Moreover, we noticed that a supportive environment that allowed student teachers to express critical and self-critical views was fruitful for the emergence of reflection. However, the study has certain limitations. First, the coding of the data was made by only one researcher, which diminishes the reliability of the study. Another limitation is study's restriction to one student teacher group with one teacher educator. Moments of reflection emerged regularly in the discussions, but it is uncertain how much these depended on the participants, especially the proficient teacher. Authentic assessment is an arrangement that has the potential to evoke reflection, but based on the present study, we cannot generalise the findings regarding its efficiency. Cognitively, the students did not seem to need the teacher educator's participation to reach the level of reflection. One could argue that the organisation of authentic assessment, especially the explicit criteria, made the students more independent from the teacher and provided them with the tools to evaluate their own performance and make arguments about others' performances. However, when it comes to a safe environment, the teacher educator's role remains uncertain. It is possible that students might have built a safe environment themselves, but in this case, the teacher educator systematically created it. In addition, it is important to bear in mind that reflection is affected by many contextual factors (Bergh, Ros, and Beijaard 2015) and there may have been other factors besides the safe environment and criticism that encouraged the reflection but that were not identified.

Assessment is often treated as a somewhat separate element of learning that, with effort, is aligned with learning objectives and attached to the learning process. The current study presents an alternative possibility in which the learning unit is *based* on assessment. Here, building a learning unit on assessment consisted of (1) choosing theory relevant to students' professional performance and introducing it to students, (2) illustrating dimensions of theory with concrete criteria, (3) guiding students videotape their authentic performance, evaluate it with the criteria, and choose a clip which presents a good example of an element of

theory, and (4) watching video clips of students' performances and discussing those based on criteria and with teacher's guidance. Our way of planning assessment diverges from the proposition of Villarroel et al. (2018); in their model, feedback comes at the end of planning. We suggest that feedback dialogues could function as the main method of learning. The key element is research-based criteria that make theory, assessment, and learning discussions inseparable so that assessment paradoxically becomes inconspicuous. In this model, peer assessment resembles colleagues' discussions more than schoolwork, which is in line with the idea of authentic assessment (Villarroel et al. 2018). Authentic assessment seems like a good fit in teacher education, especially with teaching practice, but we propose that building learning units on authentic assessment could also be beneficial in other contexts.

Acknowledgements

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

Notes on contributors

Laura Ketonen is a university teacher at the Department of Teacher Education and a researcher at the Finnish Institute for Educational Research, University of Jyväskylä. Her research concerns assessment and feedback in the contexts of primary and secondary education, and higher education. ORCID: https://orcid.org/0000-0002-2821-0179

Minna Körkkö is a post-doctoral researcher at the Faculty of Education and Psychology, University of Oulu. Her research interests include teacher professional development and teacher competence as well as student teachers' reflection. ORCID: https://orcid.org/0000-0002-4248-1434

Sanni Pöysä is a post-doctoral researcher at the Department of Teacher Education, University of Jyväskylä. Her research foci include teacher-student interaction and student engagement as well as teacher and student well-being. ORCID: https://orcid.org/0000-0002-5059-9459

ORCID

Laura Ketonen   http://orcid.org/0000-0002-2821-0179
Minna Körkkö   http://orcid.org/0000-0002-4248-1434
Sanni Pöysä   http://orcid.org/0000-0002-5059-9459

References

Afdal, H. W., and K. Spernes. 2018. "Designing and Redesigning Research-Based Teacher Education." *Teaching and Teacher Education* 74 (August): 215–228. https://doi.org/10.1016/j.tate.2018.05.011.

Beauchamp, C. 2015. "Reflection in Teacher Education: Issues Emerging from a Review of Current Literature." *Reflective Practice* 16 (1): 123–141. https://doi.org/10.1080/14623943.2014.982525.

Bergh, L. V. D., A. Ros, and D. Beijaard. 2015. "Teacher Learning in the Context of a Continuing Professional Development Programme: A Case Study." *Teaching and Teacher Education* 47 (April): 142–150. https://doi.org/10.1016/j.tate.2015.01.002.

Biggs, J. B., and C. So-Kum Tang. 2011. *Teaching for Quality Learning at University: What the Student Does*. 4th edition ed. Maidenhead, England New York: SRHE and Open University Press Imprint.

Black, P., and D. Wiliam. 2009. "Developing the Theory of Formative Assessment." *Educational Assessment, Evaluation and Accountability(formerly: Journal of Personnel Evaluation in Education)* 21 (1): 5. https://doi.org/10.1007/s11092-008-9068-5.

Carless, D., and D. Boud. 2018. "The Development of Student Feedback Literacy: Enabling Uptake of Feedback." *Assessment & Evaluation in Higher Education* 43 (8): 1315–1325. https://doi.org/10.1080/02602938.2018.1463354 .

Christ, T., P. Arya, and M. Ming Chiu. 2014. "Teachers' Reports of Learning and Application to Pedagogy Based on Engagement in Collaborative Peer Video Analysis." *Teaching Education* 25 (4): 349–374. https://doi.org/10.1080/10476210.2014.920001.

Chye, S., M. Zhou, C. Koh, and W. Chia Liu. 2019. "Using E-Portfolios to Facilitate Reflection: Insights from an Activity Theoretical Analysis." *Teaching and Teacher Education* 85 (October): 24–35. https://doi.org/10.1016/j.tate.2019.06.002 .

Danielowich, R. M. 2014. "Shifting the Reflective Focus: Encouraging Student Teacher Learning in Video-Framed and Peer-Sharing Contexts." *Teachers & Teaching* 20 (3): 264–288. https://doi.org/10.1080/13540602.2013.848522.

Darling-Hammond, L., and J. Snyder. 2000. "Authentic Assessment of Teaching in Context." *Teaching and Teacher Education* 16 (5): 523–545. https://doi.org/10.1016/S0742-051X0000015-9.

Dawson, P., D. Carless, and P. Pui Wah Lee. 2021. "Authentic Feedback: Supporting Learners to Engage in Disciplinary Feedback Practices." *Assessment & Evaluation in Higher Education* 46 (2): 286–296. https://doi.org/10.1080/02602938.2020.1769022.

Dewey, J. 1933. *How We Think: A Restatement of the Relation of Reflective Thinking to the Educative Process*. Boston: D. C. Heath and company.

Henriette, L., and R. F. Poell. 2016. "On Critical Reflection: A Review of Mezirow's Theory and Its Operationalization." *Human Resource Development Review* 15 (1): 3–28. https://doi.org/10.1177/1534484315622735 .

Jay, J. K., and K. L. Johnson. 2002. "Capturing Complexity: A Typology of Reflective Practice for Teacher Education." *Teaching and Teacher Education* 18 (1): 73–85. https://doi.org/10.1016/S0742-051X0100051-8 .

Kember, D., D. Y. P. Leung, A. Jones, A. Y. Loke, J. McKay, K. Sinclair, H. Tse, et al. 2000. "Development of a Questionnaire to Measure the Level of Reflective Thinking." *Assessment & Evaluation in Higher Education* 25 (4): 381–395. https://doi.org/10.1080/713611442 .

Ketonen, L., M. Hähkiöniemi, P. Nieminen, and J. Viiri. 2020. "Pathways Through Peer Assessment: Implementing Peer Assessment in a Lower Secondary Physics Classroom." *International Journal of Science and Mathematics Education* 18 (8): 1465–1484. https://doi.org/10.1007/s10763-019-10030-3.

Ketonen, L., S. Lehesvuori, S. Pöysä, and M.-K. Lerkkanen. 2021. "Contradictions and Reflection: Teachers Views About Interactive Formative Assessment." *NMI Bulletin* 31 (2): 41–60. https://bulletin.nmi.fi/2021/06/22/ristiriitoja-ja-reflektiota-opettajien-taydennyskoulutuksessa-esittamia-nakemyksia-vuorovaikutuksessa-tapahtuvan-arvioinnin-haasteista-ratkaisuista-ja-itse-koulutuksesta/ .[In Finnish]

Kolb, D. A. 1984. *Experiential Learning: Experience as the Source of Learning and Development*. Englewood Cliffs, N.J: Prentice Hall.

Körkkö, M. 2020. "Beneath the Surface : Developing Video-Based Reflective Practice in the Primary School Teacher Education Programme." PhD diss., University of Lapland. https://lauda.ulapland.fi/handle/10024/64109.

Körkkö, M. 2021. "Towards Meaningful Reflection and a Holistic Approach: Creating a Reflection Framework in Teacher Education." *Scandinavian Journal of Educational Research* 65 (2): 258–275. https://doi.org/10.1080/00313831.2019.1676306 .

Körkkö, M., O. Kyrö-Ämmälä, and T. Turunen. 2016. "Professional Development Through Reflection in Teacher Education." *Teaching and Teacher Education* 55 (April): 198–206. https://doi.org/10.1016/j.tate.2016.01.014.

Korthagen, F. 2017. "Inconvenient Truths About Teacher Learning: Towards Professional Development 3.0." *Teachers & Teaching* 23 (4): 387–405. https://doi.org/10.1080/13540602.2016.1211523 .

Korthagen, F., and A. Vasalos. 2005. "Levels in Reflection: Core Reflection as a Means to Enhance Professional Growth." *Teachers & Teaching* 11 (1): 47–71. https://doi.org/10.1080/1354060042000337093.

Lauriala, A. 2013. "Chapter 27 Changes in Research Paradigms and Their Impact on Teachers and Teacher Education: A Finnish Case." In *From Teacher Thinking to Teachers and Teaching: The Evolution of a Research Community*, edited by Cheryl J. Craig, Paulien C. Meijer, and J. Broeckmans, Vol. 19. 569–595 . Advances in Research on Teaching, Emerald Group Publishing Limited. https://doi.org/10.1108/S1479-3687(2013)0000019030

Levin, B., and Y. He. 2008. "Investigating the Content and Sources of Teacher Candidates' Personal Practical Theories (PPTs)." *Journal of Teacher Education* 59 (1): 55–68. https://doi.org/10.1177/0022487107310749 .

Matsumoto-Royo, K., and M. Soledad Ramírez-Montoya. 2021. "Core Practices in Practice-Based Teacher Education: A Systematic Literature Review of Its Teaching and Assessment Process." *Studies in Educational Evaluation* 70 (September): 101047. https://doi.org/10.1016/j.stueduc.2021.101047 .

McGarr, O., and O. McCormack. 2014. "Reflecting to Conform? Exploring Irish Student Teachers' Discourses in Reflective Practice." *The Journal of Educational Research* 107 (4): 267–280. https://doi.org/10.1080/00220671.2013.807489.

Mezirow, J. 1991. *Transformative Dimensions of Adult Learning.* 1st edition ed. San Francisco: Jossey-Bass.

Nieminen, J. H., and L. Tuohilampi. 2020. "'Finally Studying for Myself' – Examining Student Agency in Summative and Formative Self-Assessment Models." *Assessment & Evaluation in Higher Education* 45 (7): 1031–1045. https://doi.org/10.1080/02602938.2020.1720595 .

Onwuegbuzie, A. J., and N. L. Leech. 2007. "Validity and Qualitative Research: An Oxymoron?" *Quality & Quantity* 41 (2): 233–249. https://doi.org/10.1007/s11135-006-9000-3.

Paksuniemi, M., P. Keskitalo, S.-M. Frangou, and M. Körkkö. 2021. "Pre-Service Teachers' Experiences of Dialogical and Reflective Supervision Through Digital Technology." *International Journal of Technology in Education and Science* 5 (3): 463–485. https://doi.org/10.46328/ijtes.243.

Peltier, J. W., A. Hay, and W. Drago. 2005. "The Reflective Learning Continuum: Reflecting on Reflection." *Journal of Marketing Education* 27 (3): 250–263. https://doi.org/10.1177/0273475305279657 .

Pöysä, S., E. Pakarinen, A. Kajamies, and M.-K. Lerkkanen. 2021. "VOPA Procedure Supporting teachers' Assessment and Interaction Skills." *NMI Bulletin* 31 (2): 81–93. [In Finnish]

Pöysä, S., E. Pakarinen, L. Ketonen, S. Lehesvuori, and M.-K. Lerkkanen. 2021. "VOPA-Procedure's Phased Descriptions of Learning interaction."." *Jyväskylä University Digital Repository*. https://jyx.jyu.fi/handle/123456789/75754. [In Finnish]

Rodríguez-Gómez, G., and I.-S. María Soledad. 2015. "Assessment as Learning and Empowerment: Towards Sustainable Learning in Higher Education." In *Sustainable Learning in Higher Education: Developing Competencies for the Global Marketplace*, edited by M. Peris-Ortiz and J. M. M. Lindahl, 1–20. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-10804-9_1 . Innovation, Technology, and Knowledge Management.

Schön, D. A. 1983. *The Reflective Practitioner: How Professionals Think in Action*. London: Temple Smith.

Stockero, S. L., R. L. Rupnow, and A. E. Pascoe. 2017. "Learning to Notice Important Student Mathematical Thinking in Complex Classroom Interactions." *Teaching and Teacher Education* 63 (April): 384–395. https://doi.org/10.1016/j.tate.2017.01.006.

Tai, J., R. Ajjawi, D. Boud, P. Dawson, and E. Panadero. 2018. "Developing Evaluative Judgement: Enabling Students to Make Decisions About the Quality of Work." *Higher Education* 76 (3): 467–481. https://doi.org/10.1007/s10734-017-0220-3.

Toom, A., H. Kynäslahti, L. Krokfors, R. Jyrhämä, R. Byman, K. Stenberg, K. Maaranen, and P. Kansanen. 2010. "Experiences of a Research-Based Approach to Teacher Education: Suggestions for Future Policies." *European Journal of Education* 45 (2): 331–344. https://doi.org/10.1111/j.1465-3435.2010.01432.x .

Valli, L. 1997. "Listening to Other Voices: A Description of Teacher Reflection in the United States." *Peabody Journal of Education* 72 (1): 67–88. https://doi.org/10.1207/s15327930pje7201_4.

Van Manen, M. 1977. "Linking Ways of Knowing with Ways of Being Practical." *Curriculum Inquiry* 6 (3): 205–228. doi:10.1080/03626784.1977.11075533.

Villarroel, V., S. Bloxham, D. Bruna, C. Bruna, and C. Herrera-Seda. 2018. "Authentic Assessment: Creating a Blueprint for Course Design." *Assessment & Evaluation in Higher Education* 43 (5): 840–854. https://doi.org/10.1080/02602938.2017.1412396.

Winstone, N. E., R. A. Nash, J. Rowntree, and M. Parker. 2017. "'It'd Be Useful, but I Wouldn't Use It': Barriers to University Students' Feedback Seeking and Recipience." *Studies in Higher Education* 42 (11): 2026–2041. https://doi.org/10.1080/03075079.2015.1130032.

Appendix

Classroom organization - Video clip 1

reflection

noticing

Classroom organization - Video clip 2

reflection

noticing

Classroom organization - Video clip 3

reflection

noticing

Classroom organization - Video clip 4

reflection

noticing

Feedback – Video clip 1

Reflection

Noticing

Feedback – Video clip 2

Reflection 2

Reflection

Noticing

Feedback – Video clip 3

Reflection

Noticing

Feedback – Video clip 4

Reflection

Noticing

Feedback – Video clip 5

Reflection

Noticing