JYX
JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

**This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.**

**Author(s):** Ketonen, Laura; Nieminen, Juuso Henrik

**Title:** Supporting student teachers' reflection through assessment : The case of reflective podcasts

**Year:** 2023

**Version:** Published version

**Copyright:** © 2023 the Authors

**Rights:** CC BY 4.0

**Rights url:** https://creativecommons.org/licenses/by/4.0/

**Please cite the original version:**

Ketonen, L., & Nieminen, J. H. (2023). Supporting student teachers' reflection through assessment : The case of reflective podcasts. Teaching and Teacher Education, 124, Article 104039. https://doi.org/10.1016/j.tate.2023.104039

**The University of Jyväskylä's Rights Retention Strategy (RRS) enables accepted versions (final drafts) to be made openly available through self-archiving. For more information, see: https://www.jyu.fi/en/rrs**

Teaching and Teacher Education 124 (2023) 104039

Contents lists available at ScienceDirect

Teaching and Teacher Education

journal homepage: www.elsevier.com/locate/tate

Research paper

# Supporting student teachers' reflection through assessment: The case of reflective podcasts

Laura Ketonen [a, *], Juuso Henrik Nieminen [b, c]

[a] Department of Teacher Education, University of Jyväskylä, Jyväskylä, Finland
[b] Faculty of Education, The University of Hong Kong, Hong Kong SAR
[c] Centre for Research in Assessment and Digital Learning (CRADLE), Deakin University, Melbourne, Australia

ARTICLE INFO

*Article history:*
Received 1 July 2022
Received in revised form
21 December 2022
Accepted 16 January 2023
Available online xxx

*Keywords:*
Reflection
Assessment
Teacher education
Authentic assessment
Digital assessment

ABSTRACT

Reflection is a key learning objective of teacher education, yet little is known about the ways in which it could be supported through assessment. In this case study (N = 14), we examine the potential of a novel assessment task—reflective podcasting—in supporting first-year student teachers' reflection. We analyzed student group podcasts and corresponding student reflections and found that the task engendered reflection. Three theory-driven dimensions were constructed to emphasize the unique nature of podcasts for student reflection: experientiality, authenticity, and communality. We encourage factoring these dimensions into assessment designs that seek to foster student reflection.

© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

## 1. Introduction

Reflection is considered a central process of teacher learning and has become one of the most important goals of teacher education (Collin & Karsenti, 2011; Körkkö et al., 2016; Korthagen, 2017). Reflection is crucial for learning new educational theories and concepts as well as developing teaching practice based on them. While an extensive body of literature has sought to examine how teaching practices in pre-service teacher education could promote student reflection, the role of *assessment* in this equation has received less attention. In the present study, our goal is to determine whether and how an assessment task in the format of a podcast promotes student reflection.

Assessment serves multiple purposes in teacher education (Bloxham, 2008), for example, the summative purpose of accountability, which ensures that graduating teachers possess the skills they are supposed to master before entering the profession. As reflection is an important learning objective of teacher education, rigorous methods are needed to summatively assess whether students have achieved it. At the same time, assessment should support learning, for example, through formative assessment practices, such as self- and peer assessment (e.g., Lynch et al., 2012). Besides serving summative and formative functions, assessment should also prepare students to "meet their own future learning needs" (Boud, 2000, p. 151) and serve their lifelong learning (Boud & Soler, 2016). We argue that assessment that supports reflection also supports student teachers' immediate and lifelong learning because of the centrality of reflective skills in teachers' learning (Korthagen, 2017; Ottesen, 2007).

However, research has shown that supporting reflection through assessment is not a straightforward process (Lalor et al., 2015). For example, Minott (2008) found that reflective journals only enabled a few students to reach the level of critical reflection, that is, examining and disrupting the social and political presumptions of schooling. Importantly, the wider context of assessment and grading that overshadows teacher education warrants further discussion as assessment and grading policies are not always designed with student reflection in mind. Researchers have warned that an overemphasis on summative assessment of reflection may decrease student teachers' ability and willingness to reflect (Meeus et al., 2009), shift their approach from being learning-oriented to performance-oriented (Ramsey, 2010; Roberts, 2016), and limit their free expression by creating pressure

* Corresponding author. PO Box 35, FI-40014, University of Jyväskylä.
*E-mail addresses:* laura.k.ketonen@jyu.fi (L. Ketonen), juuso@hku.hk (J.H. Nieminen).

https://doi.org/10.1016/j.tate.2023.104039
0742-051X/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

to please the evaluator (Chan et al., 2021). According to Bloxham (2008), the multiple purposes of assessment often contradict each other, and the summative function of accountability may create barriers to achieving goals such as "critical reflection," for example, by reducing students' willingness to take risks in assessment (Chan & Lee, 2021).

In the present study, we revisit the pertinent question by Lalor et al. (2015, p. 46): "[C]an a process of professional reflective practice be initiated through an assessment format that challenges trainee-teacher's existing beliefs and perceptions and encourages reflection on, in and for future action?" What we add to the existing literature is an exploration of whether the *podcast format*, as a digitally mediated assessment task, could promote student reflection in pre-service teacher education and how. Podcasts are audio-based digital media files that often take the form of storytelling or a dialogue (McHugh, 2016). Audio files themselves are not a novel idea in education, but podcasts offer an intriguing platform for assessment, given their timely and authentic format (cf. Nieminen, Bearman, & Ajjawi, 2022). They offer potential affordances for learning within the context of "the digital world," especially since we now live with technology rather than only making use of it (Bearman et al., 2022). The educational benefits of podcasts have variously been documented in the higher education and school teacher practices literature (e.g., Besser et al., 2021; Gunderson & Cumming, 2022) as well as in some studies on pre-service teacher education (Carson et al., 2021; Forbes & Khoo, 2015; Souter & Muir, 2008). Following Carson et al. (2021), we propose that podcasts offer the potential not only for imminent learning benefits but also for students' longer-term reflection. They may enable students to take part in reflective, dialogic spaces in ways that the more traditional form of reflective writing cannot (Chan et al., 2021; Fullana et al., 2016). In this study, we focus on this potential through an empirical analysis.

### 1.1. Reflection in teacher education

The concept of reflection originates from Dewey (1933), who defined it as the intentional and determined examination of beliefs, knowledge, and practice based on existing information and understanding and considered it an important educational goal. While some researchers view reflection as challenging and that it reconstructs prior knowledge and conceptions, others relate it to the active and purposeful exploration of experiences (Chan & Lee, 2021). Both viewpoints are valuable in teacher education, which requires both the disruption and rebuilding of one's beliefs about education and the examination of teaching practice experiences. In our study, we rely on Mezirow (1991) conceptualization of reflection, which belongs to the former category.

Most of the theorizations about reflection and teacher learning are action-bound. For example, Schön (1987) identified three forms of reflection that all relate to practice: reflection-on-action, which occurs when planning lessons; reflection-in-action, which takes place during teaching; and reflection-on-action, which happens retrospectively after teaching. Korthagen's (2017) cyclical model of teacher reflection (ALACT) begins with action, continues with memorizing this action, becoming aware of its central aspects, and creating alternative ways to act, finally returning to action with a new trial. The connection between theory and practice is central to action-bound models.

In the present study, our approach to reflection is theory-bound. The intention is not to overlook the value of action-bound reflection for teacher learning but to represent the role of reflection in the early stages of teacher education when students rarely have personal experience of teaching. We draw on Körkkö et al. (2016) to argue that the main aim of reflection in students' first year of teacher education is constructing a research-based knowledge base, which includes an examination of one's own experiences and becoming aware of one's own beliefs and their origins. Many well-known reflection models used in teacher education (e.g., Korthagen, 2017; Valli, 1997) are not built for this purpose as they relate to teaching experiences. Reflection is content- and context-specific (Korthagen, 2017). Even though individuals can develop their ability to reflect (Körkkö et al., 2016), it depends on environmental factors, such as social aspects (Stoll et al., 2006) and the content of the reflection (Lee, 2005).

In this study, we use Mezirow (1991) theory-based approach to reflection. Mezirow comes from the field of adult learning, and his conceptualizations of reflection are based on his transformative learning theory. He considers it essential for learning that individuals become aware of their experiences and backgrounds and the influence that these factors have on their thinking. For Mezirow, reflection is a tool for questioning one's reasoning and is aimed at emancipation and a change in perspective. Such an emancipatory approach to reflection is relevant in teacher education, where students join following years of educational experience, although in the role of pupil (Körkkö et al., 2016). During these years, they adopt views about teaching and learning, many of which are likely to be false. Moreover, as knowledge about learning and teaching continues to increase, it is vital that teacher education does not simply reproduce teachers but, rather, ensures that each teacher generation understands the state of the art. Accordingly, Mezirow (1991) approach to reflection is particularly valuable as learning requires a disruption of prior beliefs. Moreover, the interconnection of the cognitive and social aspects of Mezirow's conceptualization of reflection serves our research aims. An additional rationale for the choice of Mesirow's conceptualization is its applicability. Scientists have built on his work and developed various instruments that determine levels of reflectivity, including definitions of non-reflection (Lundgren & Poell, 2016).

We employ the four-level reflection continuum of Peltier et al. (2005) to conceptualize reflectivity and its depth. This model builds on Mezirow's conceptualization of reflection (Lundgren & Poell, 2016). On the lowest level of the continuum, learning is unreflective, superficial, and based on memorization (see Table 5). The second-lowest level resembles book learning, where the aim is to understand new theories and concepts. The third level is called *reflection*; it entails the connection of new theories and concepts to one's own experience and knowledge, the critical examination of one's preconceptions, and consideration of alternative ways of thinking and acting. The fourth and highest level of the continuum is *intensive reflection*, which includes meta-level observations, that is, the examination of the origins of one's thoughts, perceptions, and actions and alterations of fundamental personal beliefs. One of the strengths of Peltier et al. (2005) model for our study is that it separates reflection and non-reflection, unlike some otherwise informative models that have been used in teacher education (Jay & Johnson, 2002; Korthagen, 2017; Valli, 1997).

## 2. Research questions

In this study, we empirically examine the potential of podcasting for first-year teacher students' reflection as a form of digitally mediated assessment (cf. Bearman et al., 2022; Nieminen et al., 2022). Using the case study approach on one university course (N = 14, in four groups), we introduce a podcast assessment task aimed at both determining the educational content of the course and promoting students' reflection. First, we analyze the level and topics of reflection in the student group podcasts. By topics, we refer to the themes that the students associate with the content of the book (social psychology), not the content itself.

**RQ1).** What kind of reflection did the student groups express in the podcasts?

a) Employing the four-level framework of Peltier et al. (2005), what level of reflection did the student groups express in the podcasts?
b) What topics were addressed in the student group reflections?

RQ1 was addressed through a *positivist-minded analysis* in which we aimed to analyze the levels of reflection as carefully as possible, that is, a "small q" qualitative analysis. However, we also wanted to understand the social and cultural factors influencing the students' reflection in the podcasts. Thus, we conducted another analysis, which employed a *social interpretivist epistemology*. With this twofold view, we ended up with two radically different qualitative analyses, akin to a mixed-methods study. We asked the students to take part in follow-up group conversations and individual interviews concerning the podcasts, which we examined through a "big Q″ qualitative analysis following Braun and Clarke's (2021) framework for reflexive thematic analysis.

**RQ2).** As the students discuss the benefits and hindrances of podcasting for their reflection, what kinds of socio-cultural dimensions are highlighted?

RQ2 considers the fact that first-year students are likely unable to critically analyze their own reflection skills or recognize the factors influencing them. The socio-culturally oriented analysis provided us with theoretical tools to reframe the students' imminent reactions to the task.

## 3. Method

### 3.1. Context and participants

The case study took place in a teacher education program at the University of Jyväskylä, Finland. Finnish teacher education can be described as research-based, and student teachers learn both academic and research skills (Toom & Husu, 2021). All student teachers graduate with a master's degree. The teacher educators at the University of Jyväskylä have high autonomy in both teaching and assessment. There are no national guidelines for assessment, and even the department's guidelines are flexible. A great deal of teaching occurs in small groups, where teacher educators can adapt the teaching and assessment based on their pedagogical judgment. While the study units are graded, the grades have little if any effect on the students' future lives and careers.

The "case" in our study is one university course titled Interaction and Collaboration. We present this course as a *typical case* in preservice teacher education as it consists of ordinary components: lectures, small-group meetings, and mandatory literature. The study participants were 14 first-year teacher education students. All the students participated in a program for primary school teachers, and some of them were also studying to become secondary school teachers and earn a master's degree in mathematics or science. At the University of Jyväskylä, student teachers are assigned to "home groups," and in their first year, they complete most of their educational studies in this group (for more information, see Juutilainen et al., 2018). All 14 students belonged to the same home group, and this was their first unit of educational studies together. The first researcher was responsible for teaching the educational studies units during the year, and she knew the students well. This study unit (Interaction and Collaboration, 5 ECTS credits) included five 90-min lectures, seven 90-min small-group meetings, and independent work. The course curriculum is described in more detail in Table 1.

At the beginning of the course, the teacher instructed the student group to decide, within the limits of the curriculum, what kinds of learning tasks they wanted to perform during the course and how they wanted them to be assessed (pass/fail vs. grade, what kind of feedback). After a few rounds of planning and discussion, the students decided on assessment based on two artifacts: a learning diary and a podcast. The learning diary would be written individually and graded, and the podcast would be made in small groups and assessed with a pass/fail grade. However, the students wanted the podcast to increase their grade if its level was higher than that of the learning diary, a request accepted by the teacher. Based on the students' plans, the teacher wrote instructions for the podcast task (Table 2).

One rationale for the choice of podcast was the students' endeavor to cope with the mandatory literature. The largest individual piece of literature was a 300-page book on social psychology. Instead of every student having to read the book individually, the sample was divided into four groups for the purposes of reading and processing their share of the book and creating a podcast about it for the other students. The students listened to the other groups' podcasts and provided peer feedback. Besides informing the groups about the quality of their podcast, the feedback functioned as a control for listening.

As shown in the course curriculum, reflection was not covered as a topic in the course literature, neither were reflection skills an explicit learning objective of the course. However, the teacher and students had multiple discussions about reflection in relation to the assessment of the course tasks. The assessment criteria (Table 3) were based on reflection and demonstrated its importance in learning educational knowledge. The teacher specifically emphasized that even though the understanding of the concepts and theories might have been a successful strategy in secondary school, the situation was different in teacher studies. The same criteria were applied to different courses and assessment tasks. Despite the wording of the criteria, the object of the assessment was a collectively created podcast, not an individual exercise.

### 3.2. Data collection

As is typical of case studies, we constructed multiple data sets to gather a holistic understanding of our research topic (Yin, 2018). The data sets consisted of the podcasts, reflective group discussions about the podcasts, and interviews (Table 4). With a few guiding questions, the groups discussed the podcast experience about six weeks after completing the study unit and recorded the discussion themselves. The students were asked to discuss their experience of the podcast from the perspectives of learning, emotions, and motivation. They were also asked to compare the podcast to other assessment tasks. The discussions were not graded. Thereafter, the students who volunteered (N = 3) for the semi-structured interviews were asked more about the podcast task and their preferences and previous experiences with assessment. The second author conducted the interviews from an outsider position in order to reduce the power imbalance between the interviewer and the interviewees. We anticipated that the students would speak more freely, particularly about potentially critical views, as the interviewer was not their teacher.

### 3.3. Analysis

#### 3.3.1. RQ1) What kind of reflection did the student groups express in the podcasts?

To answer RQ1, we performed a deductive qualitative content analysis (Schreier, 2012). Here, the students' utterances were the unit of analysis, which we categorized as one of the four types of

**Table 1**
Details of the course curriculum.

| |
|---|
| **Course name**: Interaction and Collaboration (5 ECTS credits)<br>**Grading scale:** 0–5<br>**Description of the content:**<br>- Interaction theories and their application<br>- Ethics and cultural diversity in interaction<br>- Observation and investigation of interaction<br>- Own interaction skills<br>**Learning outcomes:**<br>After the course, students will be able to:<br><br>- Observe interaction in diverse contexts, analyze them from the chosen perspectives, and examine the emotions emerging within themselves and others<br>- Listen to others and express themselves clearly, taking cultural variety into account<br>- Examine individuals as part of the group and the formation of group dynamics and communality<br>- Apply their understanding and interaction skills in diverse situations (e.g., feedback and conflict situations) and interact in an ethically sustainable way<br>**Mandatory literature:**<br>1. Suoninen, E., Pirttilä-Backman, A.-M., Lahikainen, A. R., & Ahokas, M. (2013). *Arjen sosiaalipsykologia* [Social psychology]. Sanoma Pro.<br>2. An English journal article about cultural diversity [of the teacher's choice] |

**Table 2**
Podcast instructions for students.

| |
|---|
| General instructions<br>- The length of the podcast is 20–30 min.<br>- Study the *Social Psychology* book with your group. Each group reads a fourth of the book thoughtfully and creates a podcast for other groups to listen to.<br>*Instructions for reading*<br>- Select a section of the book that you will all read independently, and discuss among yourselves. Choose how many sections you want to share reading.<br>- Once you have finished reading, select the most central content for the podcast.<br>*Instructions for podcast*<br>- The podcast should cover theory as well as your considerations and practical examples of the content. Try to make the podcast enjoyable to listen to.<br>- Consider how you will record the podcast. Mp3 is the recommended format.<br>- Return the podcast on time to ensure enough time for listening.<br>*Feedback about podcasts*<br>- After listening to the podcasts of others, provide feedback.<br>- The emphasis should be on encouraging feedback. Describe what you liked about the podcast. In addition, you can suggest how it could be improved.<br>*Assessment of podcasts*<br>- Assessment of the podcasts is pass/fail. The learning diaries are graded.<br>- However, it the grade of the podcast is higher based on the same assessment criteria, the grade of learning diary will be raised by one unit. |

**Table 3**
Assessment criteria for the course (an applied five-step version taken from Peltier et al., 2005).

| 1 Describing | 2 Understanding | 3 Analyzing | 4 Reflection | 5 Intensive Reflection |
|---|---|---|---|---|
| The student concentrates on memorization and repetition of concepts and themes and does not connect the task to the big picture. The student shows little engagement and a surface approach to learning. | The student focuses on understanding concepts and themes but does not connect them to their own experiences or prior knowledge. The learning resembles book learning. | The student connects themes and concepts to their own experiences, describes their own conceptions and ideas about them, and provides concrete examples. The student shows active engagement. | The student connects themes and concepts to their own experiences and prior knowledge. The student challenges their prior conceptions and identifies areas for improvement. The student shows active engagement and has a deep approach to learning. | The student reflects on why they think and act as they do and where their conceptions of the concepts and themes are derived from. The student may change their personal beliefs and ways of thinking as a result of learning. |

**Table 4**
The data sets of the study and how they relate to the research questions.

| RQ | Data set | Description of the data set |
|---|---|---|
| RQ1 | Student podcasts | Four podcasts (27:27, 34:35, 35:52, and 24:33 min) |
| RQ2 | Student group discussions about the podcasts | Four group discussions after the course (13:43, 17:29, 16:07, and 10:36 min) |
| RQ2 | Individual student interviews | Three interviews after the course (34:33, 29:10, and 30:08 min) |

reflection using the framework of Peltier et al. (2005; see Table 5). If the utterance contained several levels of reflection, it was coded under the highest category. All the material was initially coded by the first author. In order to make sense of the codes in their local contexts, we coded each of the podcasts separately rather than analyzing a whole pool of codes. This was followed by categorizing the codes. Various measures were used to ensure the trustworthiness of the analysis. First, the authors began the analytical process by discussing the framework and analyzing 30 units together. This was followed by several research meetings through which we calibrated our understanding of the analytical framework. Second, the data set (430 coded units) was analyzed by both authors, after

**Table 5**
The non-reflection–reflection continuum of Peltier et al. (2005).

| Non-reflection/Surface Learning <——————————> Reflection/Deep Learning | | | |
|---|---|---|---|
| Habitual Action (P1) | Understanding (P2) | Reflection (P3) | Intensive Reflection (P4) |
| Minimal thought and engagement, correlated with a surface approach to learning—specific tasks are treated as unrelated activities, memorization is emphasized, embodying 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 only needs to comprehend the materials read. 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. It shows active and conscious engagement, characteristics commonly associated with a deep approach to learning. | Intensive reflection is at the highest level of the reflective learning hierarchy, and learners become aware of why they think, perceive, or act as they do. Learners might alter or even reject firmly held beliefs and ways of thinking. Intensive reflection is thus seen as involving a change in personal beliefs. |

which they held a meeting to discuss their agreements and discrepancies and calculated the weighted κ for the podcast data set to determine intercoder reliability based on the recommendation by Feng (2014). Intercoder reliability for the levels of reflection was excellent (weighted kappa = .94). The high kappa value was due to the intense and dialogic preparatory phase before the whole data set was analyzed separately by the two authors.

Next, we continued to categorize the extracts containing reflection or intensive reflection. To determine the topics of reflection corresponding with the student group podcasts, we inductively categorized the coded units. To capture the students' meaning-making as authentically as possible, we combined *descriptive* and in vivo *coding* (Saldaña, 2021); that is, we used the students' own words and sayings as codes (in vivo coding) and supplemented them with our interpretation (descriptive coding).

#### 3.3.2. RQ2) As the students discuss the benefits and hindrances of podcasting for their reflection, what kinds of socio-cultural dimensions are highlighted?

To address our second research question, we followed Braun and Clarke's (2021) methodology of reflexive thematic analysis (RTA) to analyze the follow-up group discussions and interviews. RQ2 explicitly drew on an interpretivist paradigm to unpack the socio-cultural elements underlying and influencing reflection and provided us with tools to analyze the potentially surface-level reflections of the first-year students about such socio-cultural processes.

We started the process by carefully familiarizing ourselves with the data set. We then coded the *group discussions* about the podcasts following the process of descriptive and in vivo coding of RQ1, resulting in 121 codes. Examples of the in vivo codes included "I felt that I had to read it really carefully" and "Familiar people sharing examples," and the corresponding descriptive codes were "group pressure" and "intimacy," respectively. Based on the codes, we started to produce initial themes collaboratively. In this phase, the process was inductive, and the first set of nine themes was solely data-driven. This inductive starting point allowed us to remain sensitive to the contextual elements of the data. As we reviewed the initial themes, we reflected on the data-driven themes with existing literature on assessment and reflection. In this way, we reiterated the initial themes by starting a dialogue between ourselves (authors), the data set, and the research literature. At this stage of the analysis, we engaged in reading to understand our data; we read about various topics but continuously returned to assessment as learning, authentic assessment, and sustainable assessment. The literature guiding our data-informed final themes was not determined before but during the process. After coming to a mutual agreement on the interpretation, we analyzed the *student interviews* through a similar process to that explained above. The interviews offered us a way to triangulate our findings. Their role was affirmative; they strengthened, deepened, and contested our thematization, leading to the finalization of the themes. Finally, following Braun and Clarke (2021), our final report was not separate from the analysis but, rather, an integral part of it. The final report fully represents our interpretation, and the writing process itself crystallized our mutual sense-making of the data sets. We recognize that it is not conventional to report "findings" and "earlier literature" in tandem as we do, but only in this way can we shed light on which aspects were discovered through the data sets in a theory-informed, dialogic manner.

Given the interpretivist epistemology of RQ2, we relied on Lincoln and Guba's (1985) concepts of credibility, transferability, and confirmability. First, *credibility* was ensured by inviting the students to a member-checking session to discuss our findings. *Transferability* was ensured through a thick description of our data and the context of Finnish teacher education. *Confirmability* was supported by the dialogic, reflexive analysis process, which drew on our collaborative meaning-making.

## 4. Results

### 4.1. RQ1) What kind of reflection did the student groups express in the podcasts?

In the podcasts, we identified all the reflection levels in Peltier et al. (2005) (Table 5). The mandatory literature framed the students' conversations, so the utterances at all levels were related to the content of social psychology. This is logical since the aim of the podcast was to help other students learn the content. We also identified an additional level (P0) that was typically related to facilitating podcasts, including comments to listeners, such as "Welcome to our podcast," and other discussants, such as "Do you want to tell us about the book; what kind of experience was it?"

The utterances at P1 were the least frequent. Here, the students explained the concepts and theory by quoting the book or their own notes. For example, in the following excerpt, a student introduced a concept to the discussion:

> S6: We also have social representations. They are systems of values, ideas, and practices that have two functions. First, they create a structure that enables people to orientate to the material and social world and makes the world controllable. (P1)

The concept of social representations is a difficult one. Therefore, to capture it accurately, the students quoted the book or their notes instead of explaining it in their own words.

Level P2 was the most common, and it was aimed at understanding. The utterances at this level included students explaining and discussing the concepts and theories, often by giving general examples:

S12: Prejudices are a subgroup of attitudes, and they mean attitudes toward certain people. For example, you presuppose someone to be of a certain type even without having met them or knowing anything about them. (P2)

One-fourth of the students' speaking turns in the podcasts contained reflection or intensive reflection (Table 6). Levels P0, P1, and P2 were important components of the podcasts, but as they did not rise to the level of reflection, which is the focus of this study, they were precluded from further analysis. The students' reflections concerned five main topics: shared experiences of university life, individual experiences of university life, topical phenomena, experiences outside university life, and consideration and elaboration of theory (Table 7). These were all identified in several group conversations. However, only 10 utterances contained intensive reflection (P4), nine of which came from one group.

*Consideration and elaboration of theory* means reflecting on the theory from different viewpoints and sometimes even challenging it. The theories concerned, for example, relationships, group formation, and early interaction. The following excerpt exemplifies relationships:

S12: It [the book] mentioned that a friendship can be a basis for a long intimate relationship, that they are intertwined. And the book said that the only thing separating the traditional intimate relationship and a friendship is that in a friendship, your best friend can have other friends, but in a traditional intimate relationship, your partner seldom has other companions. But we can criticize that this book was written several years ago, and therefore it may not have been able to consider the perspective of open relationships. (P3)

*Experiences outside university life* entailed students' experiences and views about their childhood, everyday life, friends and family, hobbies, feelings, and values. These were typically used as illustrations of concepts and theories to help listeners internalize them. The following excerpt relates to a section in the book about attachment theory and the need for attachment in stressful situations:

S2: Besides this, the need for attachment becomes obvious in stressful situations. I have noticed that even though I sometimes want to be alone, I need others and seek support when I'm stressed. Have you experienced the same? (P3)

*Topical phenomena* included conversations about the COVID-19 pandemic and its influences, interaction in digital environments, and social media. For example, in the following extract, a student elaborates on a section in the book about the influence of technology on children's development:

S13: And social media, if we return to the bullying theme, it creates a new environment that connects the school and social media. It is scary how it [bullying] can continue at home and everywhere. (P3)

**Table 6**
The levels of reflection on the podcasts.

| | N | % |
|---|---|---|
| **P0** | 90 | 20.9 |
| **P1** | 28 | 6.5 |
| **P2** | 206 | 47.9 |
| **P3** | 96 | 22.3 |
| **P4** | 10 | 2.3 |

**Table 7**
Topics of reflection on podcasts.

| Topic | N (P3) | N (P4) |
|---|---|---|
| Consideration and elaboration of theory | 26 | 0 |
| Experiences outside university life | 25 | 0 |
| Topical phenomena | 22 | 0 |
| Shared experiences of university life | 13 | 10 |
| Individual experiences of university life | 10 | 0 |

The only topic to elicit intensive reflection concerned *shared experiences of university life*: the organization of the students' studies, their home group's qualities, and their ways of interacting in the home group. As the name of the topic suggests, intensive reflection emerged when the content of the book matched the students' common experiences. For example, in the following excerpt, a student analyzes how they act as a group in their studies and why. The *how* question is a level P3 reflection, and *why* belongs to level P4:

S3: This brings to mind that when we discuss with a whole group, we hesitate to talk. But when we're divided into smaller groups, intense discussion begins. It may be because, in a smaller group, you feel obligated to talk and share your opinion. But also, it is easier to share these opinions to the whole group [after discussing in a small group], as you can say that it's your group's opinion. If you answer something by yourself, you may fear that you might be wrong and that no one will support your opinion. (P4)

*Individual experiences of university life* related to students' new life situations: moving to a new city, learning to know new people, and studying at the university. The students reflected on various social phenomena through their experiences:

S6: Yep, they support each other well, the different social networks. It would be harsh to study at the university if there were no friends. It would hardly support my well-being. (P3)

### *4.2. RQ2) As the students discuss the benefits and hindrances of podcasting for their reflection, what kinds of socio-cultural dimensions are highlighted?*

The RTA process resulted in three theory-driven themes concerning the dimensions of reflective podcasting and emphasizing the uniqueness of this format in the students' accounts: experientiality, authenticity, and communality. Each of these themes was described by the students as both hindering and promoting reflection. In this section, these themes are presented separately only for the sake of analysis and clarity. In actuality, they did not represent "separate" themes but were tightly woven together. Thus, they do not "cover" the data set but offer a full and precise set of interwoven interpretations.

#### *4.2.1. Experientiality*

In higher education, assessment is predominantly presented as a systematic and technical matter. The grading systems identifiable in most higher education contexts render assessment an apparatus of measurement and comparison. The idea of outcome-based assessment views competence as "subdivided into separate measurable, stable traits" (Hodges, 2013, p. 565) and assessment as a way to determine precisely whether such measurable outcomes have been achieved. Such processes reflect the basis of assessment

in the measurement paradigm. This view has been challenged by emphasizing the relational, socio-material, and embodied aspects of assessment. Experientiality presents the importance of the experiences provided by the assessment task and the significance of sharing experiences for reflection. Experientiality reframes assessment as a complex, non-reductive endeavor, not as something that is "done to" students but as "a means for students to embody and affectively experience knowledge, rather than seeing it as an impossible to attain standard external to them" (Gravett et al., 2021, p. 12). This theme emphasizes that students were not simply asked to reflect on assessment. Rather, they were asked to *experience* podcasting, including the creation and choice of the task, revealing the relational and emotional aspects of knowing and understanding educational theory (see Fawns et al., 2021).

*4.2.1.1. How did experientiality promote reflection?.* The task provided the students with a *lived experience* of podcasting. The students largely referred to the assessment task with emotional adjectives, such as "exciting," "different," and "real." A sense of novelty and excitement was mentioned by many of them as a factor enhancing reflection:

> S1: It felt novel. I had never done a podcast before, so recording one made me feel very nervous. But in the end, I had a very good feeling about it because our group was so good. The conversations we had during the process were excellent. (Interview)[1]

Experientiality resonates with Dewey, who considered experience as a starting point of reflection and reflection as a process of giving meaning to experience (Rodgers, 2002). The podcast provided the students with an experience of assessment that was divided between various digital and non-digital spaces. This was exemplified by S2. The following excerpt shows how the final podcast presented only the final product of an ongoing process of reflection divided among physical spaces, WhatsApp, and temporal aspects:

> S2: The teacher gave us freedom to decide how we wanted the course to be assessed. I was nervous about whether we could really make it. It felt difficult to start, but we got our act together and did it. It went all right. I enjoyed making it, and we booked a working space and worked there. And we also had discussions in our WhatsApp group; we did some planning there too. (Interview)

The podcast format enabled the students to share personal and emotional experiences as part of the process of learning. The experiences were interwoven into the knowledge provided by the textbook, allowing reflection to emerge. For example, S3 noted in the group discussion that "I feel that I remember those [theories]. It probably helped that you had to think about how they were connected to you. Had we only replicated the theory from the book, it might be that no one would remember anything." Tying the educational theories to part of one's personal experiences offered possibilities for deeper reflection.

S2 pondered whether the podcast task would have been able to promote reflection in disciplines in which knowledge is considered less relational than in teacher education. This finding emphasizes the disciplinary context of teacher education as a potential venue for podcasting due to the specific idea of knowledge as something relational, that is, something that can be lived through (see Nieminen & Lahdenperä, 2021). S2 spoke about their major (mathematics) in terms of its "right or wrong epistemology," which does not allow personal and embodied experiences to be interwoven into learning:

> S2: Well, math is like, it is right or wrong. You cannot connect your own experiences to it. You cannot refer to your emotions, experiences, or anything. (Interview)

The data sets for RQ2 unveiled a process that could not be illuminated by the original podcasts: how listening to other groups' podcasts promoted the students' learning and reflection. Listening to the podcasts not only presented educational theories in a more accessible digital format but also enabled the students to learn from each other's personal reflections. For example, S4 mentioned that by personally knowing the students in the recordings, they could "reach the other person's mental landscape compared to reading the text and giving feedback about it." Similarly, the students indicated that the podcasts enabled them to meaningfully relate to the experiences of others. Relatable reflections enabled further reflection, as explained by S5:

> S5: When I normally read a book, I miss over half of it. But now, when I listened to others' [podcasts] and them sharing their own experiences, I was able to make associations with them.

*4.2.1.2. How did experientiality hinder reflection?.* The students described their earlier assessment histories hindering their reflection. First, they were not accustomed to reflection in *assessment*. This was seen most clearly in the individual interviews, in which the students compared the podcast task with their earlier experiences of assessment in high school. S1 explained that inexperience in such reflective assessment tasks made the podcast task "difficult." S3 noted that their assessment in high school had been almost completely based on formal essays. S1 further stated that their previous studying had been grade-oriented. Through such reflective university tasks, therefore, they were now forced to critically examine their grade- and test-oriented approaches to learning.

A second feature of the students' assessment histories that hindered their reflection was that they were not accustomed to complex assessment tasks that had abstract learning goals, such as reflection. S2 did not enjoy the lack of clarity and the *messiness* of podcasting and hoped for more objective learning measurements, probably the kind experienced in their prior school years:

> S2: I would change it [the assessment] to be more straightforward—that you were told more explicitly what the grades correspond to. I feel that it is sometimes vague. And the teacher could maybe have more control, I think. (Interview)

*4.2.2. Authenticity*

The second theme concerned the digitally mediated authenticity of podcasting. Authentic assessment often refers to assessment that uses "the same competencies, or combinations of knowledge, skills, and attitudes, that they need to apply in the criterion situation in professional life" (Gulikers et al., 2004, p. 69). We expand this view by relying on the conceptualization of digitally mediated authentic assessment (Nieminen et al., 2022) as something that connects to the digital world we live in rather than only considering digital tools for assessment. Authentic assessment was then seen to prepare the students for their future profession as

[1] The extracts from the interviews have been marked; if not specifically stated otherwise, the extracts are from group discussions.

educators in digital knowledge societies (Bearman et al., 2022).

*4.2.2.1. How did authenticity promote reflection?.* The podcast, representing a popular form of media, provided an authentic digital form of assessment. Podcasts inherently include personal reflections—rarely do they aim to simply *educate* their listener (Carson et al., 2021). As such, they enabled an authentic, digitally mediated way to learn about and reflect on educational theory.

Authenticity was seen in the way in which the students oriented their reflection to the world around them. First, the digital format enabled and even required them to show and represent their authentic selves. S7 argued that the podcast "necessitated improvisation rather than a tight script." Moreover, the process of creating the podcast engendered authentic conversations in digitally mediated ways:

> S6: With essays, you read the literature and write, and you don't really think about what you are doing, whereas with podcasts, you have genuine conversations about the topic with the group.

The emotions that the students shared in the podcasts elicited a sense of vulnerability. This relates to the idea of digital vulnerability in authentic assessment (Lynch & Sargent, 2020). Similar to Lynch and Sargent's video narrative task, the podcast task "allowed the students in this study to be 'vulnerable', 'expressive' and the embodied connections encouraged psychological emotions as an affective process" (p. 8). Moreover, the students could authentically include the world around them as a meaningful part of the task by reflecting on the educational theories in light of their current experiences, for example, the COVID-19 pandemic or their newly begun university student life:

> S13: Familiar people sharing examples was good, especially when there were connections to my own life. For example, our home group was used as an example. I still remember those examples, but I recall very few from the book.

A crucial aspect of providing an authentic experience of reflection was the multimodality of the podcast. This was something that the students addressed in each group discussion, stating that they listened to the podcasts while driving, jogging, or doing dishes. S7 even raised multimodality as a matter of accessibility and shared how their dyslexia restricted their written reflection while the spoken, recorded form of the podcast offered a more accessible way to reflect:

> S7: I have a rather severe form of dyslexia, and therefore, writing is difficult for me. I would rather discuss the topic and show my skills better that way.

*4.2.2.2. How did authenticity hinder reflection?.* Vulnerability is always risky in authentic assessment and digital spaces (Lynch & Sargent, 2020), and this was the case in the podcast task. While the sense of digitally mediated vulnerability and meaningfulness separated the podcast from "inauthentic" practices, such as exams and essays, these aspects also engendered stress and excitement about the task. For example, S1 noted that the planning phase for the podcast was "stressful" and that the group was "under pressure," thus providing a deep reflection on a recorded tape did not feel completely natural. Other students indicated that the feelings of excitement and stress might have hindered reflection:

> S8: I guess that I could have shown my skills better by writing because I'm nervous in such situations. I freeze up and cannot share my thoughts clearly.

> S9: I had frustrated feelings in the beginning and nervousness about how we would succeed. Before recording, I felt that we were kind of in a panic (laughs). We almost did not have the courage to start recording. But of course, the memory also raises good feelings because the conversation was brilliant. But we did not notice that until we finished the recording.

The students discussed the balance between spontaneity and planned reflection in the podcast format, which showed how this "authentic" form of media also erected barriers to recorded reflection. This *authenticity* of the podcast format and *inauthenticity* in the conversations between the students were seen as restricting the students' reflection alongside the abovementioned emotions of nervousness and stress.

S7: We had a sort of script, but afterwards, we felt that we might have been able to improvise the conversation. If it hadn't been planned, we could have made it more personal, without such a strict [script]. But, on the other hand, that was our first, at least my first podcast ever, which caused nervousness—having a script created safety.

S1 mentioned that their group shared meaningful discussions about the topics outside the recording situation but that these *authentic* reflections were never caught on tape:

> S1: My honest opinion is that our podcast failed in spontaneity because we were so nervous. We had such good conversations outside the recording, but the planning made the podcast strict. (Interview)

*4.2.3. Communality*

The final theme concerned the shift from the individualistic underpinnings of assessment to a *communal* epistemology. In higher education, the structures of assessment predominantly reflect the individualist ideologies of performativity, competition, and accountability (Nieminen & Lahdenperä, 2021; Gravett et al., 2021). Higher education certifies individual students through its grading mechanisms—not groups of students. This presents profound challenges regarding any attempt to promote collective approaches to assessment: assessment is commonly conceptualized based on the individualistic idea of psychometric measurement (Hodges, 2013). From a sociopolitical point of view, assessment in higher education is "a crucial technology through which students are positioned in the market and through which they are encouraged to see themselves as entrepreneurial citizens" (Gravett et al., 2021, p. 11). Previous literature on digitally mediated assessment has neglected the communal aspects of assessment within and beyond the university setting (Nieminen et al., 2022). It might be that the "dialogic spaces" that Carson et al. (2021) analyzed might have been particularly tricky to uphold in podcasting that is *assessed* and *graded*. The theme of communality shifts the unit of analysis from individual students to communities of students, rendering the process of reflecting on educational theory a communal endeavor rather than an individualistic one.

*4.2.3.1. How did communality promote reflection?.* While group work itself was not new to the students, communal approaches to assessment were. All three interviewees described how assessment in high school rarely included collaborative or communal dimensions. The podcast provided an embodied experience of

shared, communal knowledge as the educational theories were not *known* by individuals but by the group, and so was reflection. This aligns with earlier research findings on how deeper reflection in relation to assessment might require a lived experience rather than simply a reflective conversation about a topic (Nieminen & Lahdenperä, 2021). Social communication and interaction were largely seen as reflection enablers:

> S2: Talking alone and recording would not be a good idea. This kind of taskmust be done in groups. That brings the sense of conversation, especially in our group; we got deep into the topic and shared our own experiences. (Interview)

We emphasize the idea of communality over simple collaboration as the podcast was more than the sum of its parts. The students did not simply divide the work among themselves; instead, they engaged in a communal discussion. This allowed multiple voices and experiences to be heard, which disrupted the individualist epistemologies of assessment. While the educational theories in the course book could arguably be understood in *incorrect* ways, the knowledge inside the book was internalized through deep reflection enabled by communal epistemologies. "Different kinds of ideas and experiences" (S11) and "different viewpoints" (S9) were needed to reflect on the material. It was not the individual students who were doing the learning and being assessed but, rather, the groups, as shown in the following excerpt:

> S3: What made it specifically intriguing was perhaps dividing the book so that every group could read part of it. It did not create pressure to read the whole book and internalize it, but you were able to trust that the other groups had picked up the necessary information about their theme, and you would learn it when you listened to it. Reading only one piece raised your motivation—once you knew that you wouldn't have to read the whole book.

Importantly, the final products—the podcasts—were not rendered meaningless after the students completed the assessment task. Instead, the cycle of learning and reflection continued as the students listened to the podcasts recorded by others and learned from their lived experiences and authentic reflections:

> S12: I liked that, as we had different groups and different concepts, there were examples that I could not have come up with. That somehow augmented my understanding better than just reading the book alone and staying with my own thoughts because they often go only one way.

*4.2.3.2. How did communality hinder reflection?*. Some students raised issues about collaborative work overall, such as having to co-regulate the whole process with all the deadlines and arrangements. For example, H3 noted that, overall, the students had a great deal of group work in their studies, and therefore, another group task compounded the stress. Another student reflected on the upsides and downsides of communal approaches to assessment:

> S1: There are pros and cons [in group work]. … I like working in a group, but on the other hand, things like scheduling and other decision-making, those have always had their own challenges. (Interview)

## 5. Discussion

In the present study, we analyzed a podcast task to determine whether and how it might influence student reflection. Based on an in-depth qualitative analysis, we showed that the podcast engendered reflection (Table 6) on several topics related to educational concepts and theories (Table 7), and a theory-oriented thematic analysis revealed three dimensions that both supported and hindered reflection: experientiality, authenticity, and communality. Overall, the findings shed light on how assessment could support reflection in the context of teacher education and, thus, serve students' current and future learning needs.

Only a quarter of the podcast reflections were intensive, and it emerged when the students connected educational concepts and theories to their shared experiences, pondering why they felt or acted as a group in the way they did. Two specific features of the reflective podcast task enabled these moments. The first was the collaborative assessment task, and the second was content relating to the students' shared experiences (of phenomena of social psychology). This finding suggests that teacher education should reconsider the use of typical assessment tasks that do not invite students to co-reflect: portfolios, exams, learning diaries, and essays. A comparison of the reflection rates in this study to those of other studies was not informative as the consistency of the reflection criteria between the studies was low, and the variance between the outcomes, even those related to teacher education, was large (Dyment & O'Connell, 2011; Lundgren & Poell, 2016). Moreover, unlike most existing research, we assessed the reflection of groups rather than individuals. Identifying reflection on podcasts was notable since it demonstrated that reflection could be promoted and assessed with other artifacts besides commonly used reflective texts.

While we have already discussed the findings in the context of the assessment literature in the results section, we continue with additional considerations. *Experientiality* of assessment derived from the possibility of sharing experiences within the assessment task and the experienced complexity and messiness of podcasting. The first-year students were accustomed to test-driven assessment, where the teacher selects the assessment tasks and assesses student performance individually. In this case, the students were guided to select the assessment tasks themselves and choose the impact they would have on their grade. Similar to earlier studies, allowing students to have an influence on their assignments and, thus, challenging traditional teacher and student roles was a positive experience for the students (Juutilainen et al., 2018). However, the lack of traditional rules on assessment was sometimes bewildering, and many students wanted greater clarity, even though they recognized that sharing their experiences supported their reflection and that accurate assessment of their learning might not be possible.

*Authenticity* influenced the students' experiences of the task. In teacher education, reflective journals, portfolios, and essays are commonly used in promoting and assessing student reflection. While such artifacts are often created and shared in digital environments, their authenticity can be questioned because they are not necessarily formats that students use in their free time or will use in their future work as teachers. Moreover, as students do not consider these formats as important tools for supporting their reflection (Chye et al., 2019), we suggest that reflection could be supported by diverse multimodal assessment tasks. Different tasks have distinct features, and as noted by Carson et al. (2021), in the podcast format, the dialogic space and accessibility were the

central elements promoting the authenticity of the task and reflection. Rather than only considering podcasting as a potential format for promoting reflection, we challenge readers to consider which methods and formats would fit the idea of digital authenticity in their own contexts (see Nieminen et al., 2022).

*Communality* was described as a principal dimension supporting reflection. This is in line with the findings of Collin and Karsenti (2011), who underlined the collective dimension of reflection, arguing that verbal interactions were at the center of student teachers' reflection. Mutual discussion and trust are commonly regarded as supporting reflection (Körkkö et al., 2016; Mezirow, 1991), and therefore, aligning learning and assessment (i.e., tasks in small groups) with the learning objective (i.e., reflection) supports learning.

As assessment policies have not necessarily been created with students' reflectivity in mind, research should, besides examining how assessment could support reflection, examine how to avoid assessment becoming *a hindrance* to reflection. Our findings are in line with Bloxham's (2008) notion regarding the contradictory purposes of assessment. Therefore, the development of new assessment practices is only one aspect in efforts aimed at supporting reflection; equally important are considerations about how current policies and practices hinder students' reflection and how these effects could be mitigated.

### 5.1. Implications for practice

Perhaps surprisingly, we do not suggest that podcasting would necessarily promote reflection in all or even most teacher education contexts. Instead, through this specific podcast task, we were able to identify the three enabling and hindering themes relating to reflection. We believe that these themes are beneficial for reflective assessment practice in the digital world (Bearman et al., 2022), and we challenge teacher educators to consider each of them when designing assessments.

The idea of *experientiality* aligns with Bloxham's (2008) suggestion that assessment in teacher education does not need to avoid unclarity and complexity but, rather, cherish it (see Fawns et al., 2021; Hodges, 2013). As students gain embodied experiences of assessments that bring forth the messy, entangled, and complex aspects of reflection, they might also gather resources for deeper future reflection (Nieminen & Lahdenperä, 2021). *Authenticity* reminds us that reflection tasks that seem meaningful and authentic might be more beneficial in supporting reflection than inauthentic practices. In other times and places, other formats besides podcasts might seem more "authentic" to students and should thus be considered. *Communality* has been integral to ideas such as shared reflection, but this ideal is rarely fully implemented in assessments. Based on our findings, we encourage teachers to promote communal reflection in ways that explicitly disrupt the individualistic underpinnings of assessment and grading and might promote critical reflection on such assessment systems in teacher education. Our findings also suggest that communal assessment tasks might be especially useful if the students have shared experiences of the content to be learned. As discussed earlier, all of these dimensions may also hinder reflection, and they should be contextually considered in situational implementations of "assessment for reflection."

Finally, we note that the various ideas herein can be implemented in practice in the early phases of teacher education programs. We acknowledge the challenges of reshaping assessment due to students' assessment histories and higher education assessment policies. However, because of the fixed nature of assessment traditions, the disruption of assessment practices could be particularly effective in causing a *disorienting dilemma* that triggers reflection and leads to perspective change (Mezirow, 1991). We argue that assessment should be employed at the start of teacher studies so as to challenge student teachers' thinking and preconceptions of education, supporting the examination of their beliefs, a change in their perspective, and emancipation. It may be that this starting phase is exactly when assessment should be harnessed for the purpose of deep reflection, giving students critical tools to proceed in their future studies and assessment.

### 5.2. Limitations and implications for future studies

Our study has some limitations. In seeking to empirically understand the interconnections of assessment and reflection, we relied on an in-depth, small-scale case study. Such a design has obvious limitations regarding the sample size and amount of data. While the reflection framework of Peltier et al. (2005) was useful for making sense of the podcast data set, more nuanced frameworks (e.g., Korthagen, 2017) might allow future studies to unpack the nuances of multimodal forms of reflection. The first author was the teacher of the course, which might have affected the students' responses in the group discussions and interviews, despite our strategies for preventing such issues (see the methods section). However, we note that the students shared their critical and even openly negative views of podcasting, which might be symptomatic of the informal, non-hierarchical nature of Finnish teacher education. Our qualitative study did not aim for generalizability; however, we discussed the transferability of the findings, which we sought to enhance through thick descriptions of the context and data (Lincoln & Guba, 1985). Nevertheless, we acknowledge that the situated nature of our study reduces the transferability of our findings to teacher education contexts beyond Finland. In addition, our study only offers a snapshot of one university course. Future studies could explore the temporal aspects of reflection through longitudinal research designs (see Chan et al., 2021). Finally, while we highlighted the experientiality, authenticity, and communality of "assessment for reflection," future studies could capture the messiness and complexity of reflection using more diverse methods. For example, our interviews only captured a tiny portion of how reflection was constructed in informal spaces, such as in the students' WhatsApp groups. Innovative digital data sets might be able to capture these processes more fully in future research.

## 6. Conclusions

Reflection is a core competence in teacher education. Thus, it needs to be included more coherently in assessment design. Reflection skills take time to develop (Körkkö et al., 2016). Therefore, rather than making single attempts at integrating assessment and reflection, the assessment of reflection should be programmatic. We suggest that assessment has particular potential for promoting student reflection and call for further research to unpack how the dimensions of experientiality, authenticity, and communality play out in assessment designs in various teacher education contexts.

## Data availability

The data that has been used is confidential.

## References

Bearman, M., Nieminen, J. H., & Ajjawi, R. (2022). Designing assessment in a digital world: an organising framework. *Assessment & Evaluation in Higher Education*. https://doi.org/10.1080/02602938.2022.2069674. Advance online publication.

Besser, E. D., Blackwell, L. E., & Saenz, M. (2021). Engaging students through educational podcasting: Three stories of implementation. In *Technology,*

*knowledge and Learning*. Advance online publication. https://doi.org/10.1007/s10758-021-09503-8.
Bloxham, S. (2008). Assessment in teacher education: Stakeholder conflict and its resolution. *Practitioner Research in Higher Education, 2*(1), 13–21.
Boud, D. (2000). Sustainable assessment: Rethinking assessment for the learning society. *Studies in Continuing Education, 22*(2), 151–167.
Boud, D., & Soler, R. (2016). Sustainable assessment revisited. *Assessment & Evaluation in Higher Education, 41*(3), 400–413.
Braun, V., & Clarke, V. (2021). One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative Research in Psychology, 18*(3), 328–352.
Carson, L., Hontvedt, M., & Lund, A. (2021). Student teacher podcasting: Agency and change. *Learning, Culture and Social Interaction, 29*, Article 100514.
Chan, C. K. Y., & Lee, K. K. W. (2021). Reflection literacy: A multilevel perspective on the challenges of using reflections in higher education through a comprehensive literature review. *Educational Research Review, 32*, Article 100376.
Chan, C. K. Y., Wong, H. Y. H., & Luo, J. (2021). An exploratory study on assessing reflective writing from teachers' perspectives. *Higher Education Research and Development, 40*(4), 706–720.
Chye, S., Zhou, M., Koh, C., & Liu, W. C. (2019). Using e-portfolios to facilitate reflection: Insights from an activity theoretical analysis. *Teaching and Teacher Education, 85*, 24–35.
Collin, S., & Karsenti, T. (2011). The collective dimension of reflective practice: The how and why. *Reflective Practice, 12*(4), 569–581.
Dewey, J. (1933). *How we think: A restatement of the relation of reflective thinking to the educative process*. D. C. Heath and Company.
Dyment, J. E., & O'Connell, T. S. (2011). Assessing the quality of reflection in student journals: A review of the research. *Teaching in Higher Education, 16*(1), 81–97.
Fawns, T., Mulherin, T., Hounsell, D., & Aitken, G. (2021). Seamful learning and professional education. *Studies in Continuing Education, 43*(3), 360–376.
Feng, G. C. (2014). Intercoder reliability indices: Disuse, misuse, and abuse. *Quality and Quantity, 48*(3), 1803–1815.
Forbes, D., & Khoo, E. (2015). Voice over distance: A case of podcasting for learning in online teacher education. *Distance Education, 36*(3), 335–350.
Fullana, J., Pallisera, M., Colomer, J., Fernández Peña, R., & Pérez-Burriel, M. (2016). Reflective learning in higher education: A qualitative study on students' perceptions. *Studies in Higher Education, 41*(6), 1008–1022.
Gravett, K., Taylor, C. A., & Fairchild, N. (2021). Pedagogies of mattering: Re-Conceptualising relational pedagogies in higher education. In *Teaching in Higher Education*. Advance online publication. https://doi.org/10.1080/13562517.2021.1989580.
Gulikers, J., Bastiaens, T. J., & Kirschner, P. A. (2004). A five-dimensional framework for authentic assessment. *Educational Technology Research & Development, 52*(3), 67–86.
Gunderson, J. L., & Cumming, T. M. (2022). Podcasting in higher education as a component of universal design for learning: A systematic review of the literature. In *Innovations in Education and Teaching International*. Advance online publication. https://doi.org/10.1080/14703297.2022.2075430.
Hodges, B. (2013). Assessment in the post-psychometric era: Learning to love the subjective and collective. *Medical Teacher, 35*(7), 564–568.
Jay, J. K., & Johnson, K. L. (2002). Capturing complexity: A typology of reflective practice for teacher education. *Teaching and Teacher Education, 18*(1), 73–85.
Juutilainen, M., Metsäpelto, R.-L., & Poikkeus, A.-M. (2018). Becoming agentic teachers: Experiences of the home group approach as a resource for supporting teacher students' agency. *Teaching and Teacher Education, 76*, 116–125.
Körkkö, M., Kyrö-Ämmälä, O., & Turunen, T. (2016). Professional development through reflection in teacher education. *Teaching and Teacher Education, 55*, 198–206.
Korthagen, F. (2017). Inconvenient truths about teacher learning: Towards professional development 3.0. *Teachers and Teaching, 23*(4), 387–405.
Lalor, J., Lorenzi, F., & Justin, R. A. M. I. (2015). Developing professional competence through assessment: Constructivist and reflective practice in teacher-training. *European Journal of Educational Research*, (58), 45–66.
Lee, H.-J. (2005). Understanding and assessing preservice teachers' reflective thinking. *Teaching and Teacher Education, 21*(6), 699–715.
Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Sage Publications.
Lundgren, H., & Poell, R. F. (2016). On critical reflection: A review of Mezirow's theory and its operationalization. *Human Resource Development Review, 15*(1), 3–28.
Lynch, R., McNamara, P. M., & Seery, N. (2012). Promoting deep learning in a teacher education programme through self- and peer-assessment and feedback. *European Journal of Teacher Education, 35*(2), 179–197.
Lynch, S., & Sargent, J. (2020). Using the meaningful physical education features as a lens to view student experiences of democratic pedagogy in higher education. *Physical Education and Sport Pedagogy, 25*(6), 629–642.
McHugh, S. (2016). How podcasting is changing the audio storytelling genre. *Radio Journal: International Studies in Broadcast and Audio Media, 14*(1), 65–82.
Meeus, W., Van Petegem, P., & Engels, N. (2009). Validity and reliability of portfolio assessment in pre-service teacher education. *Assessment & Evaluation in Higher Education, 34*(4), 401–413.
Mezirow, J. (1991). *Transformative dimensions of adult learning*. Jossey-Bass.
Minott, M. A. (2008). Valli's typology of reflection and the analysis of pre-service teachers' reflective journals. *Australian Journal of Teacher Education, 33*(5), 55–65.
Nieminen, J. H., Bearman, M., & Ajjawi, R. (2022). Designing the digital in authentic assessment: Is it fit for purpose? *Assessment & Evaluation in Higher Education*. https://doi.org/10.1080/02602938.2022.2089627. Advance online publication.
Nieminen, J. H., & Lahdenperä, J. (2021). Assessment and epistemic (in)justice: How assessment produces knowledge and knowers. *Teaching in Higher Education*. https://doi.org/10.1080/13562517.2021.1973413. Advance online publication.
Ottesen, E. (2007). Reflection in teacher education. *Reflective Practice, 8*(1), 31–46.
Peltier, J. W., Hay, A., & Drago, W. (2005). The reflective learning continuum: Reflecting on reflection. *Journal of Marketing Education, 27*(3), 250–263.
Ramsey, S. J. (2010). Making thinking public: Reflection in elementary teacher education. *Reflective Practice, 11*(2), 205–216.
Roberts, P. (2016). Reflection: A renewed and practical focus for an existing problem in teacher education. *Australian Journal of Teacher Education, 41*(7), 19–35.
Rodgers, C. (2002). Defining reflection: Another look at john Dewey and reflective thinking. *Teachers College Record, 104*(4), 842–866.
Saldaña, J. (2021). *The coding manual for qualitative researchers*. Sage.
Schön, D. A. (1987). *Educating the reflective practitioner: Toward a new design for teaching and learning in the professions*. Jossey-Bass.
Schreier, M. (2012). *Qualitative content analysis in practice*. Sage.
Souter, N., & Muir, D. (2008). Podcasting, pupils and pre-service-teachers. In *European conference on educational research (ECER) 2008. From teaching to learning?*. Gothenburg, Sweden https://strathprints.strath.ac.uk/7172/.
Stoll, L., Bolam, R., McMahon, A., Wallace, M., & Thomas, S. (2006). Professional learning communities: A review of the literature. *Journal of Educational Change, 7*(4), 221–258.
Toom, A., & Husu, J. (2021). Analyzing practice, research, and accountability turns in Finnish academic teacher education. In D. Mayer (Ed.), *Teacher education policy and research: Global perspectives* (pp. 73–83). Springer.
Valli, L. (1997). Listening to other voices: A description of teacher reflection in the United States. *Peabody Journal of Education*.
Yin, R. K. (2018). *Case study research and applications: design and methods* (Sixth edition). SAGE.