Teacher Educators' Conceptions of Language Assessment Literacy in Norway

Henrik Bøhn
Department of Languages and Literature, University of South-Eastern Norway, Borre, Norway

Dina Tsagari
Department of Primary and Secondary Teacher Education, Oslo Metropolitan University, Oslo, Norway

Abstract—In order to gain a deeper understanding of what language assessment literacy (LAL) is and how it can be developed, it is relevant to investigate it from the point of view of different stakeholders (Harding & Kremmel, 2016). As for teacher LAL, this phenomenon has largely been studied from the teachers' own perspectives, leaving other stakeholders' understanding of the construct underresearched. The current study used an exploratory, qualitative research design to investigate *teacher educators' conceptions* of teacher LAL in the Norwegian educational context. Teacher educators are an influential stakeholder group in education, whose knowledge of teaching and assessment can inform our understanding of teacher LAL in important ways. Taking Taylor's (2013) LAL model as a starting point, this study explored five English teacher educators' general understanding of teacher LAL, as well as the relevance of the components of Taylor's model in particular. The study provides important empirical evidence of how teacher LAL may be understood and how Taylor's model can be further developed.

Index Terms—assessment literacy, language assessment literacy, formative assessment

I. INTRODUCTION

Language assessment literacy (LAL) may be defined as the knowledge, skills and principles needed by stakeholders in order to carry out effective assessment practices (Inbar-Lourie, 2017). This concept has attracted increasing attention over the last few years. One important reason for this is the growing importance given to assessment as a contributing factor in student learning (Black & Wiliam, 1998). Another is the empirical evidence showing that many teachers' LAL is insufficient (Jin, 2010; Lam, 2014; López Mendoza & Bernal Arandia, 2009; Vogt & Tsagari, 2014).

Several descriptions and models of LAL exist (e.g., Brindley, 2001; Fulcher, 2012; Kremmel & Harding, 2019; Taylor, 2013). One description of the construct which deserves closer attention is Taylor's (2013) LAL model.[1] This model not only distinguishes between the LAL required by various stakeholder groups, it also scales the competences needed by each group along different dimensions (see Fig. 1, below). However, Taylor makes clear that these dimensions are hypothesized, thus making them a point of reference for further discussions on how the construct can be understood. Although attempts have been made to empirically validate the dimensions (Kremmel & Harding, 2019), there are still unresolved questions relating to, for example, how they apply in specific, localized contexts, such as the one explored in the current study.

The current study takes as a starting point the teacher representation of Taylor's model. This representation puts "language pedagogy" at the heart of teacher LAL. However, the model leaves some questions unanswered, particularly regarding what some of the dimensions actually entail, and how they interrelate. Moreover, one may speculate how aspects such as disciplinary knowledge (Xu & Brown, 2016), e.g. curriculum knowledge, relate to the model. In addition, there is the question of whether the hypothesized levels of the different dimensions hold true across different contexts (Inbar-Lourie, 2013).

The present paper seeks to explore these questions, by providing empirical evidence from a Norwegian educational context. As calls have been made for the inclusion of perspectives from different stakeholders (Harding & Kremmel, 2016), the focus here will be on the views of English language teacher trainers in Norway. These are stakeholders who know the field of language teaching and assessment well, who are influential in shaping the views and practices of language teachers, and whose views on the matter have been less systematically investigated (Coombs, 2017; Villa Larenas, 2017). Their perceptions may therefore help inform our understanding of the matter in important ways.

II. LITERATURE REVIEW

The concept of LAL emerged from the more general notion of assessment literacy (AL), first introduced by Stiggins (1991), who related the term to "a basic understanding of the meaning of high- and low-quality assessment and [the

---

[1] Taylor herself did not call it a "model". However, here we will use the term in the sense of a "conceptual or mental representation of something" (OED, 2002), in line with e.g., Kremmel and Harding (2019).

ability] to apply that knowledge to various measures of student achievement" (p. 545). Descriptions and models of AL typically include both a knowledge component and a skills component (e.g., Abell & Siegel, 2011; JSCEE, 2015; Popham, 2011; Xu & Brown, 2016). The former typically relates to an understanding of assessment purposes, content and methods, whereas the latter commonly describes how to put this knowledge into practice. Some of the models also include a principles component, such as awareness of the use and impact of language tests, including issues related to ethics, fairness and professionalism (Xu & Brown, 2016; see also Davies, 2008).

Language assessment literacy (LAL) has traditionally taken as a starting point knowledge and skills related to traditional language testing, making traditional psychometric content the main focus of attention (Inbar-Lourie, 2008). In later years, however, conceptualizations of LAL have also started to include classroom-based perspectives, opening up for the possibility of including issues related to, for example, formative assessment (e.g., Brindley, 2001; Fulcher, 2012). Formative assessment can be defined as the "evaluation by brief tests used by teachers and students as aids in the learning process" (Bloom, 1969, p. 48). This may include self- and peer assessment, as well as the use of assessment by teachers to "make decisions about the next steps [to take] in instruction" (Wiliam, 2018, p. 48).

Taylor's (2013) LAL model is an example of a conceptualization which includes both knowledge, skills and principles components, and which entails both traditional testing elements and classroom-based aspects. This model is widely referred to in the language assessment literature (Baker, 2016; Deygers & Malone, 2019; Giraldo, 2018; Harding & Kremmel, 2016; Hildén & Fröjdendahl, 2018) and has been explored and applied in various contexts (Baker & Riches, 2018; Kremmel & Harding, 2019; Yan, Zhang, & Fan, 2018). It is meaningful in the way it distinguishes different skills and knowledge components along different dimensions of proficiency. Furthermore, as can be seen in Fig. 1, it distinguishes the degree of skills and knowledge needed by different stakeholders. The model is based on discussions of AL/LAL in five separate papers in a special issue of *Language Testing*, in which it appeared (Jeong, 2013; Malone, 2013; O'Loughlin, 2013; Pill & Harding, 2013; Scarino, 2013).
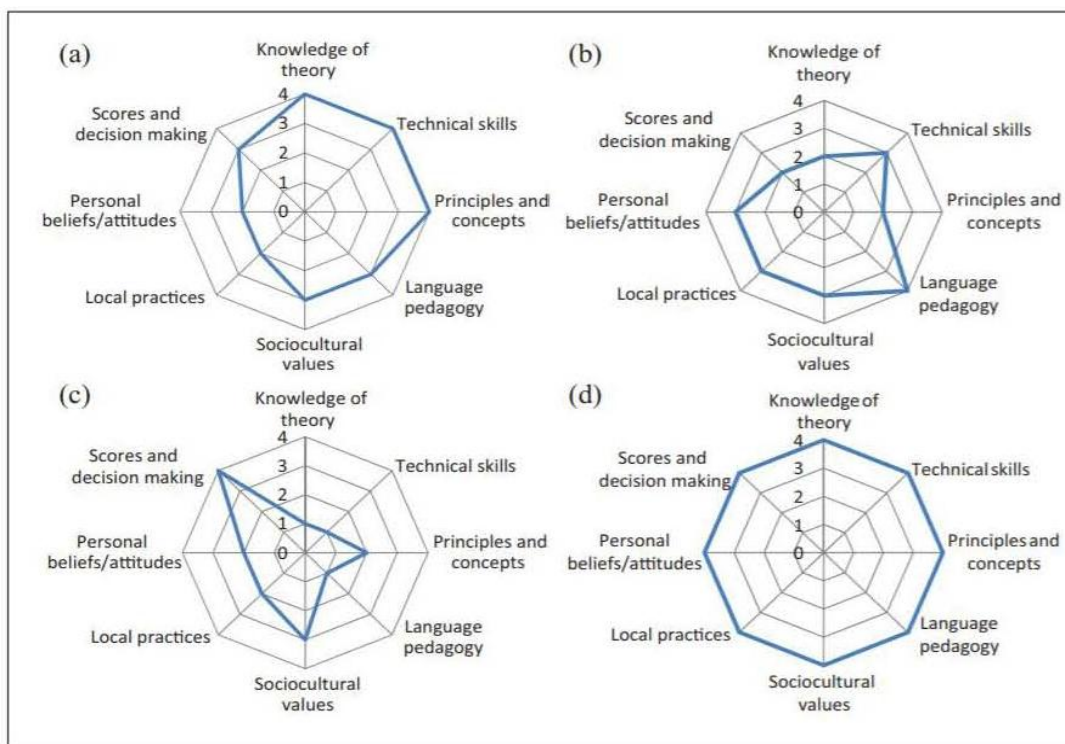

Figure 1. Taylor's (2013) LAL profiles for four different types of stakeholders
    (a)   Profile for test writers.
    (b)   Profile for classroom teachers.
    (c)   Profile for university administrators.
    (d)   Profile for professional language testers

As can be seen from profile (b) in Fig. 1, *Language pedagogy* is the component that is most significant for teachers. Conversely, *Scores and decision making, Knowledge of theory*, and *Principles and concepts* are the aspects that are least important. This may be a realization that teacher LAL is considered to be more strongly connected to learning than to testing.

From our point of view, the strength of Taylor's model lies in its emphasis on the differential levels of assessment literacy for different stakeholders. A potential challenge, however, is that it does not specify the components, or dimensions, in any great detail. Consequently, one is left to search in the other papers in the special issue of *Language Testing*, upon which the model is built, in order to understand their specific meanings. In the following, we will present

possible operationalizations of the components based on a review of the five mentioned papers in *Language Testing* and other sources directly or indirectly referred to.[2]

Starting with the component *Language pedagogy*, one may assume that this dimension is derived from Scarino's (2013) discussions on teacher LAL, which concern a number of issues related to LAL and the interrelationship between assessment, teaching and learning. However, the term "language pedagogy" is not used by Scarino. Drawing on (Shulman, 1987), Scarino points out that "pedagogical content knowledge" is central to teacher LAL. This type of competence may be defined as "knowledge that renders discipline knowledge teachable and learnable" (Scarino, 2013, p. 316). Emphasizing the relevance of formative assessment in language education, Scarino also maintains that in school contexts, "assessment cannot be separated from its relationship with the curriculum, and processes of teaching and learning" (Scarino, 2013, p. 314). What this leads us to conclude is that the component "Language pedagogy" used by Taylor (2013) includes assessment used to promote learning, as well as to the teaching of content in such a way that students will learn. The former logically concerns formative assessment, as defined by Bloom (1969) and Wiliam (2018) above, whereas the latter cogently links to disciplinary knowledge, such as language and culture (Scarino, 2013).

Moving on to *Sociocultural values*, this term is neither explained by Taylor, nor by any of the other authors of the papers in the special issue of *Language Testing*. However, Fulcher (2012), whom Taylor commends for providing a useful definition of teacher LAL, touches upon the need for teachers to have socio-cultural understanding in order to develop their assessment literacy. He includes a reference to McNamara and Roever (2006), who point to the impact that social and cultural values may have on the development and use of tests and assessments in a given context. Thus, awareness of such values can be important for teachers in order to be able to use assessments critically and appropriately.

Regarding the component *Local practices* in Taylor's model, Scarino (2013) relates this aspect to "school and classroom contexts in which [the knowledge base for teacher LAL] is practiced" (p. 315). Moreover, she seems to associate such practices with "institutional dimensions of the assessment act" (p. 322). As far as we understand, such dimensions could be curricula, laws and regulations, and locally developed assessment criteria.

The component labelled *Personal beliefs/attitudes* may logically be understood as knowledge of how one's own preconceptions, understandings and opinions may "inform [one's] conceptualizations, interpretations, judgments and decisions in assessment" (Scarino 2013, p. 109).

The component *Scores and decision making* is neither clearly defined in Taylor (2013), nor in any of the primary source papers. However, we understand "scores" as relating to teachers' knowledge of how to award grades. As for "decision making", Bachman and Palmer (2010) explain this as, for example, making decisions about whether students should be classified as "pass/fail", "certified/not certified", "beginning/intermediate/ advanced" (p. 87). Additionally, it may refer to "*formative* decisions, which are intended to help students guide their own subsequent learning, or for helping teachers modify their teaching methods and materials so as to make them more appropriate for their students' needs, interests, and capabilities" (Bachman & Palmer, 2010, p. 197, italics added). Consequently, as such practices have clear affinities with Bloom's (1969) and Wiliam's (2018) notions of formative assessment, discussed above, "decisions" seem to belong logically to the *Language pedagogy* dimension in Taylor's model. In our operationalization of this dimension, we have therefore reserved *Scores and decision making* for knowledge and skills relating to how to award grades, leaving formative decisions to the *Language pedagogy* component (see Appendix B).

Furthermore, the dimension called *Knowledge of theory* could arguably refer to both assessment and test theory, such as validity theory, and to theories of language, language learning, and communicative competence (Fulcher & Davidson, 2007). There is no direct mention of theory in Taylor's paper, but Jeong (2013), upon which Taylor partly draws, points to test theory related to "validity" and "reliability" (p. 359). However, as such theoretical knowledge could also be linked to the *Principles and concepts* dimension, as discussed below, there appears to be an overlap between these two dimensions. Nevertheless, in our study we chose to include both general test theory *and* theories of language, language learning, and communicative competence in the *Knowledge of theory* dimension (see Appendix B). We return to ways of resolving this issue in the discussion section, below.

The component *Technical skills* appears to involve skills relating to the use of appropriate methodology in test construction and use, such as "item-writing, statistics, [and] test analysis" (Davies, 2008; cited in Taylor, 2009, p. 26).

Finally, the component *Principles and concepts* seems to be taken directly from Fulcher (2012), who uses this term in his own model of teacher LAL. Fulcher points to "ethics of testing" and "test fairness" (p. 115) as examples of principles. As for "concepts", he relates this to "*theory* […] especially where statistics are introduced" (p. 124, italics added). However, as mentioned above, this causes an overlap in Taylor's model, as there is already a dimension focusing on theory. Therefore, there may be good reasons for reserving "Principles and concepts" for knowledge of ethical issues, such as test fairness, and move concepts to *Knowledge of theory* (see Appendix B).

In conclusion, Taylor's model comprises a number of interesting concepts relating to the knowledge, skills and principles needed by teachers in order to carry out good assessment practices. The lack of clear definitions admittedly complicates the operationalization of the components, particularly regarding the relationship between *Language pedagogy* and *Scores and decision making*, and between *Knowledge of theory* and *Principles and concepts*. However,

---

[2] The operationalizations (see section IV. B. below and Appendix B) will be used to investigate the extent to which the Norwegian teacher educators' find Taylor's (2013) dimensions relevant in the Norwegian context

the above explication presents some reasonable interpretations, which may serve as a starting point for empirical exploration (see Appendix B). Against this background, the present study investigated how English teacher educators in Norway perceive teacher LAL and to what extent they find the dimensions of Taylor's (2013) model relevant. The following two research questions (RQs) were therefore developed:

RQ1. What kind of LAL do Norwegian teacher educators believe that EFL teachers in Norway need in order to carry out sound assessment practices?

RQ2. How relevant do Norwegian teacher educators consider the eight dimensions in Taylor's (2013) LAL model in the Norwegian EFL context?

III. THE NORWEGIAN CONTEXT

In Norway, English is a compulsory subject taught from grade one (age six). Teaching, learning and assessment are based on a national curriculum and regulations to the Education Act. These regulations specify, among other things, that assessment should be both summative and formative. For example, the sections on formative assessment include formulations such as:

The purpose of assessment in the school subjects is to promote learning and to contribute to the joy of learning in the learning process, as well as to provide information about students' competences during the learning process and at the end of instruction […]. Continuous assessment means that students […] are to: (a) participate in the assessment of their own work and reflect on their own learning and academic development; (b) understand what they are to learn and what will be expected of them; (c) be informed of what they master; (d) receive guidance concerning how to continue improving their competence.

(Norwegian Ministry of Education and Research, 2006/2019, §3-3, 3-10, our translation)

As can be seen from this quote, teachers are required to clarify learning goals, promote feedback which is conducive to learning, and involve students in the assessment of their own knowledge and skills. These are issues which fundamentally agree with important ideas of formative assessment (Black & Wiliam, 2009; Bloom, 1969; Wiliam, 2018). By making such assessment statutory, the Norwegian educational authorities have strongly underscored that formative assessment is a fundamental principle in the Norwegian educational context. Moreover, in 2010, the government made Assessment for Learning (AfL) a prioritized area in Norwegian classrooms. To that end, a host of different seminars, conferences and online in-service courses were arranged, and the Norwegian Directorate for Education and Training developed a number of web-based resources for schools to use in their local development work in the area of AfL. The government specifically stressed that this prioritization was to be implemented on a longitudinal scale, and the programme went on from 2010 to 2018. The explicit goal of the programme was "to make school owners (i.e. local governments) [and] schools […] create an assessment culture and assessment practices where learning is the goal" (Norwegian Directorate for Education and Training, 2015, our translation).

IV. METHOD

A. *Participants*

Five teacher trainers, representing five different universities and university colleges in Norway, were recruited for the study. They were selected by means of purposive sampling (Creswell, 2013), in order to obtain variation in the sample with regard to geographical spread and types of institution (universities and university colleges). The age of the informants ranged from 48 to 64, and they all had previous experience as English teachers from the secondary school level. Four of them were holding, or about to complete, PhDs in English language education, the fifth was a trained teacher with two subsequent master's degrees in English and pedagogy. Two were trained language assessment specialists, the other three had their specializations in other fields, such as literature didactics and general English education, but they had taught language assessment modules at the tertiary level. Their teaching experience in tertiary language education ranged from eight to 23 years. The participants' identities have been anonymized in this paper, and they will be referred to as "TE1", "TE2", etc. (which are acronyms for "Teacher Educator 1", "Teacher Educator 2", and so on).

B. *Data Collection*

In order to elicit the teacher educators' conceptions of the LAL construct and their views on the relevance of Taylor's dimensions, an interview guide was developed (see Appendix A). The interview guide included both open-ended questions and closed-response items, starting with the former. The purpose of the open-ended questions at the outset was to obtain the informants' unguided reflections on what teacher LAL entails in the Norwegian context (cf. RQ1). Such a procedure could also potentially mitigate the problem of conformity in answers to closed-response items, as respondents tend to report that everything is important (Fulcher, 2012). The defining question in the opening sequence therefore read: "What kinds of skills and knowledge does an English teacher in secondary school need in order to carry out good assessments?" (cf. Appendix A).

As the study also aimed to investigate the relevance of Taylor's (2013) model in the Norwegian context, closed-response items relating to these dimensions were also included. The items, serving as operationalizations of the eight

components of the model, were formulated in line with our discussion in the literature review section above (see also Appendix B). A six-point Likert scale, going from "Not at all" to "Completely" in agreement with the statements presented, was included to measure the responses to the items.

In addition to the initial open-ended question and the subsequent closed-response items, the interview guide also comprised follow-up questions to the closed-response statements. The reason for including such questions was to give the informants the opportunity to elaborate on their answers, in case they found items to be vague or inaccurate, or wished to qualify their responses.

The interviews were carried out via telephone ($n = 4$) and Skype ($n = 1$) and lasted between 35 minutes and one and a half hour. Four out of five interviews were conducted in Norwegian, and one was carried out in English, as this was the first language of the informant. The participants were informed of the aims of the project and signed a consent form to permit us to use their data anonymously.

C. Analysis

All the interviews were transcribed by one of the authors and subsequently sent back to the teacher trainers for respondent validation (Bryman, 2016). This prompted responses from two informants concerning a couple of minor issues, such as the correct names of references given in one of the interviews. Beyond that, the informants reported being satisfied with the accuracy of the accounts.

The transcripts were analysed using the computer software QSR NVivo 12. The analysis was carried out in four cycles. In the first cycle, the transcripts were read through to get an overall impression of the material. In the second cycle, preliminary categories were created using In Vivo coding (Miles, Huberman, & Saldaña, 2014). In the third cycle, the In Vivo codes were compared internally to establish a logically consistent relationship between the categories. For example, one of the informant's reference to "disciplinary competence" was coded as a category on the superordinate level and then linked to the statement "language competence" on the subordinate level (cf. Appendix C, Informant TE1). The development of hierarchical relationships in this cycle also involved the creation of 'new', superordinate categories by the researchers, in cases where no overall description was made by the informants. One such category was "Assessment-specific competence", which was created to describe various statements, such as "be able to assess […] what should be assessed". Finally, in the fourth cycle comparisons were made across informants to arrive at an overall description of the construct of LAL. In this phase, some of the In Vivo codes were replaced by researcher-generated codes in order to achieve a more coherent description of the LAL components. An example of a researcher-generated code is the replacement of "didactic competence" by "pedagogical competence" (cf., Appendix C).

V. RESULTS

A. Results for RQ1

The results for RQ1 are based on the responses to the open-ended question "What kinds of skills and knowledge does an English teacher in secondary school need in order to carry out good assessments?". The investigation of the responses in the fourth cycle of analysis yielded the following four overall teacher LAL dimensions:

- *Disciplinary competence*
- *Assessment-specific competence*
- *Pedagogical competence*
- *Collaboration competence[3]*

The first dimension, *Disciplinary competence*, comprises aspects such as knowledge of language theory and second language acquisition, text competence, and curriculum-related know-how, including culture and literature in the English-speaking world. The teachers' own language skills may also be incorporated here, since teachers are role models to their students. Informant TE1 put it in the following way:

> Assessment competence involves, firstly, a high level of disciplinary competence… And that involves linguistic and textual competence […]. I think being a good language model is also very important, even if you can't explain [all language features]… It's very complicated. So linguistic competence also involves both knowledge and skills, right?

As can be seen from this quote, "disciplinary competence" is the component that informant TE1 first points to when explaining teacher LAL, emphasizing that a "high level" of such competence is needed. She refers to this type of proficiency as relating to language and the knowledge of texts. Similar aspects referred to by the other informants were "applied linguistics theory", "grammatical content knowledge", "knowledge about the English subject", "knowledge of theories of second language acquisition" and "knowledge of theories of communicative competence" (informants TE2, TE3, TE4 and TE5).

The second dimension, *Assessment-specific competence*, concerns aspects such as understanding the purpose of assessment, knowing what should be assessed, being able to create appropriate tasks, grading, and recognizing institutional limitations. As informant TE2 reported,

---

[3] The term "competence" has been chosen as it was used by several informants, and since it is a term widely used in the Norwegian educational context, understood by the educational authorities as involving both knowledge and skills (Norwegian Directorate for Education and Training, 2017).

you need to understand what different types of assessment are there… er… how different types of assessment have different aims or goals … er… what types of assessment do I want to use in the classroom? What type of implementational space, or what, sort of, institutional limitations are upon me? When I want to do that, what are the rules and regulations? […] And also, most importantly, I think, and it's something teachers sometimes struggle with… er… and that is… what am I doing when I'm doing this type of assessment? What is my role? And actually, what's happening?

The final remark in this quote concerning the teacher's "role" indicates that TE2 sees LAL as having a distinct metacognitive dimension. The fact that she uses the phrase "most importantly" to describe it, suggests that she finds such self-awareness to be highly significant.

The third dimension, *Pedagogical competence*, can be explained as knowledge and skills relating to how learning can be promoted, including the ability to create a good learning atmosphere, knowing how to adapt to individual students' needs, and conveying content knowledge. Communicating learning goals, assessment criteria, and the ability to give good feedback are also part of this. Informant TE5 described this dimension as:

the ability to explain, and make concrete, the assessment criteria vis-à-vis the students; to be able to show… to be able to exemplify what these… rating scales mean… at the different levels… what is needed to achieve [the mark] five, and so on.

Finally, the fourth LAL dimension, *Collaboration competence*, entails the ability of teachers to work collaboratively with colleagues in order to create a fruitful assessment culture. The ability to interact appropriately with management, parents and other stakeholders also belongs here. Informant TE1 described it as "the ability to sit down with colleagues and discuss tasks and discuss… teaching and feedback [which you receive] from them… [This involves] interaction… cooperation and … creating a common understanding".

B. Results for RQ2

RQ2, which concerned to what extent the teacher educators found the different dimensions of Taylor's (2013) model relevant for Norwegian teachers, was answered with data from the closed questions in the second phase of the interview. The overall analysis of these data showed that the teacher educators generally gave many of the items high scores. Out of a total of 125 responses 84 were scored five or six on the six-point Likert scale, indicating that the informants deemed most of the aspects mentioned important or highly important for teacher LAL. The items which received the highest scores are listed in Table 1.

TABLE 1.
THE ITEMS GIVEN THE HIGHEST SCORES BY THE INFORMANTS (TE1, TE2 ETC.) ON A SCALE FROM 1 (LOWEST) TO 6 (HIGHEST)

Dimension	Item	TE1	TE2	TE3	TE4	TE5	Mean	Median
Language pedagogy	<u>16a</u> *It is important to have knowledge of formative assessment.*	6	6	6	6	6	6.0	6.0
Local practices	<u>24</u> *It is important to have knowledge of the national curriculum.*	6	6	6	6	6	6.0	6.0
Language pedagogy	<u>12</u> *It is important to have knowledge of how assessment can promote learning.*	6	6	6	5	6	5.8	6.0
Principles and concepts	<u>28</u> *It is important to have knowledge of ethical issues.*	6	6	5	6	5	5.6	6.0
Language pedagogy	<u>30</u> *It is important to have knowledge of how to communicate assessment results*	5	6	5	6	5	5.4	5.0

As can be seen in Table 1, items 12, 16a, and 30 tap into the same dimension, i.e. the *Language pedagogy* component in Taylor's model (cf. Appendix B). Thus, the informants seem to agree with Taylor that this component is essential for teacher LAL. Moreover, the high score given to item 28, knowledge of ethics, implies that they also found the *Principles and concepts* dimension to be very important. This is contrary to Taylor's model, which suggested that this component is of less importance to teachers. As for item 24, knowledge of the national curriculum, it is worth noting that although the informants scored this item very high, the dimension to which it belongs, i.e. *Local practices*, also contains items 25 and 26 (see Appendix B), which were scored somewhat lower. Hence, this dimension was found to be slightly less important (see Fig. 2, below), and more in line with Taylor's suggestion.

As already mentioned, most items in the closed-response section received scores of 5 or 6. Very few items were given a score lower than 4. In fact, only two 2's and five 3's were awarded out of the 125 responses. The items receiving the lowest scores are shown in Table 2:

TABLE 2.
THE ITEMS GIVEN THE LOWEST SCORES BY THE INFORMANTS (TE1, TE2 ETC.) ON A SCALE FROM 1 (LOWEST) TO 6 (HIGHEST)

Dimension	Item	TE1	TE2	TE3	TE4	TE5	Mean	Median
Technical skills	15c *It is important to have knowledge of statistical measures in order to interpret results…*	2	5	3	4	4	3.6	4.0
Technical skills	15a *It is important to have knowledge of how language testing can be used appropriately…*	4	5	3	3	5	4.0	4.0
Technical skills	15b *It is important to have knowledge of how good items are created.*	6	4	3	4	5	4.4	4.0
Scoring and decision making	29 *It is important to have knowledge of how to use rating scales / scoring rubrics.*	2	6	4	4	6	4.6	5.0
Principles and concepts	31 *It is important to have knowledge of how values in society may affect assessment.*	4	6	5	5	3	4.6	5.0

Table 2 shows that issues related to Taylor's component *Technical skills* (items 15a, b, c) were found to be less important, relatively speaking, by the informants overall. This is not to say that they were found to be unimportant. In fact, there was quite a lot of variation regarding this component, as can be seen in the table. Nevertheless, the items relating to traditional testing were, on average, scored lower than the items related to the other dimensions. As for item 29 (*Scores and decision making*) and item 31 (*Sociocultural values*), they also received somewhat lower average scores. Yet, as can be seen in Table 2, this is largely due to two informants giving these items low scores. The majority of the respondents actually found these features to be important or very important.

The overall results for RQ2, which were computed by aggregating the responses to all the items for each dimension, can be seen in Fig. 2. As the teacher educators' answers were reported on a six-point Likert scale, and Taylor's original values were plotted along a five-point scale, we used a conversion table to equalize the results, using the formula 0.8x – 0.8 (IBM Support, no date). The computations were made on the basis of median scores in our findings, as they may be more relevant with ordinal data (Jamieson, 2004). The conversion table can be found in Appendix D. The results can be visualized as follows:
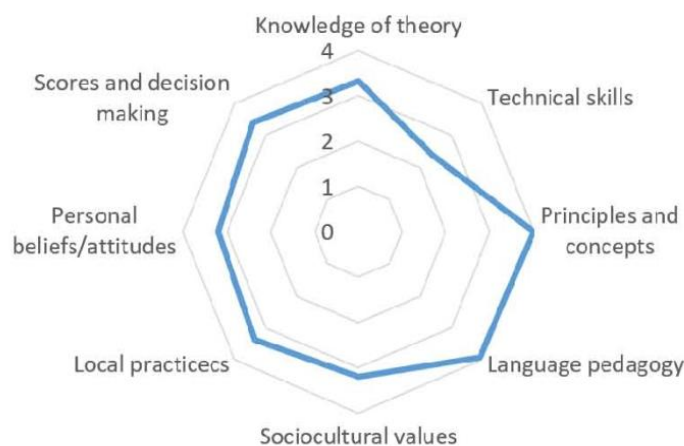

Figure 2. Teacher LAL profiles according to the Norwegian teacher educators

As can be seen in Fig. 2, the most conspicuous differences between the teacher educators' answers and Taylor's hypothesized levels are that the teacher educators found three of the dimensions considerably more important than what was predicted. This concerns:
-   *Principles and concepts* – with the value 4.00 in our findings and 2.00 in Taylor's model
-   *Scores and decision making* – with the value 3.40 in our findings and 2.00 in Taylor's model
-   *Knowledge of theory* – with the value 3.32 in our findings and 2.00 in Taylor's model

In addition, the teacher educators found the dimension *Technical skills* to be less important than what Taylor assumed (value 2.40 in our findings and 3.00 in Taylor). For the other dimensions, the differences were only marginal, although it is interesting to observe that parts of the *Local practices* dimension, i.e. knowledge of the curriculum, were found to be particularly important (cf. Table 1).

C. Follow-up Answers to the Closed-response Items

In order to explain the results arrived at above, an analysis of the follow-up answers to the closed-response items was made. Concerning the finding that the informants found *Technical skills* (e.g. knowledge of statistics) to be less important, the responses from TE1 and TE3 may shed some light on this issue. These were the two informants that scored this aspect the lowest. TE1 explained:

> TE1: *[I gave item 15c a 2,] not because I don't think it's important, but… well… it's just what you said about time being so scarce [in class].* (TE1).
> Interviewer: *Yes. Right. Er… my impression is that in Norway, there's a lot of scepticism towards testing… language testing. Do you share that impression?*
> TE1: *Yes… hmm… It's used very little, and there's very little knowledge of it, but whether there's scepticism towards it, I don't know…. It's like there's no tradition for it.*

In this exchange informant TE1 clearly states that she does not find knowledge of statistics unimportant. However, as there are so many topics to cover in the course of a teacher education programme, she thinks that this aspect could be downplayed. In addition, she thinks that there has been a Norwegian tradition for de-emphasizing language testing. This is a tradition she seems to endorse, for example by the low score she gave to item 15c (i.e., *It is important to have knowledge of statistical measures…*).

TE3 takes a somewhat different perspective, pointing to the relevance of other instruments as a means of identifying students' strengths and weaknesses in diagnostic assessment:

> *Well [you can use multiple-choice as a diagnostic tool]. But then there is the question of how you use it. Is it in order to know how many are struggling with [a grammatical problem] in the classroom? In such a case it may make sense… if the majority is struggling with* 'it *versus* there'. *But the question is: Do you need a multiple-choice test to find that out?*

In addition, TE3 reports having had more positive experiences with other forms of assessment than multiple-choice, as a young learner himself:

> *I remember I had a lot of multiple choice in school, and I often scored very poorly. But when I didn't have multiple choice, it went a lot better. Peculiar.*

Thus seen, TE3 may be sceptical for personal reasons, but also because he thinks that there are other means of eliciting responses from students, particularly in formative assessment situations, where the main purpose is to improve students' learning.

VI. DISCUSSION

The purpose of this study was two-fold: Firstly, we wanted to explore Norwegian teacher educators' conceptions of the LAL construct generally (RQ1). Secondly, we were interested in knowing to what extent they found the dimensions in Taylor's (2013) model relevant in the Norwegian context (RQ2). However, as the components of the model are not described in detail in Taylor's (2013) paper, the validity of the operationalization of the dimensions therefore rests on our interpretation of the sources upon which the model is built (see section II and Appendix B).

As for RQ1, i.e. what constitutes teacher LAL, the findings yielded four dimensions: *Disciplinary competence*, *Assessment-specific competence*, *Pedagogical competence*, and *Collaboration competence*. This is a much broader conceptualization than Taylor's model, which comprises eight dimensions (cf., Fig. 1, above). Several aspects of the dimensions in Taylor's model and our conceptualization intertwine, but there is an even stronger focus in our data on teaching-related aspects, including formative assessment and disciplinary knowledge. The finer-grained nature of Taylor's model means that a higher number of skills and knowledge are highlighted, compared to our conceptualization. Interestingly, as the results for RQ2 shows, the teacher educators actually found some of these dimensions, such as *Principles and concepts, Local practices*, and *Scores and decision making*, very relevant. One may therefore infer that a more detailed model, like Taylor's, can be quite helpful in specifying the knowledge and skills required for LAL. Kremmel and Harding's (2019) study supports this view. However, in order for the model to be applicable in educational contexts, it is essential that the dimensions are sufficiently well described and that their interrelationships are adequately clarified. Our data suggest that some components in Taylor's model need to be refined, and that the model should be supplemented with additional LAL features.

One such feature concerns the issue of disciplinary knowledge. In our analysis of RQ1, *Disciplinary competence* turned out to be an essential component of teacher LAL. This is supported by Xu & Brown (2016) and Scarino (2013). According to the teacher informants, this component involves knowledge of language and language learning, as well as curriculum-related knowledge, including culture and literature. In Taylor's model, as we interpret it, knowledge of language and language learning theories is part of the component *Knowledge of theory*. Yet, curriculum-related knowledge is not part of this dimension. This type of knowledge is included in *Local practices*. Our data suggest, however, that discipline knowledge is an important component in its own right and should therefore be singled out and highlighted as such (Xu & Brown, 2016). In order to distinguish this dimension from Taylor's *Knowledge of theory* component, we also propose that assessment-related theoretical knowledge, like for example validity theory, is put in a separate dimension labelled *Knowledge of assessment theory*. This may be seen as a modified version of Taylor's *Knowledge of theory*. Moreover, the inclusion of *Disciplinary knowledge* as a separate dimension, involving curriculum

knowledge, implies that curriculum know-how would have to be moved from Taylor's *Local practices*, where it is now included.

Another dimension which could be added to Taylor's model is *Collaboration competence*. In our analysis this component relates to the teacher's ability to contribute to the development of a shared assessment culture in his or her educational context. The idea of a shared assessment culture is supported by the Norwegian educational authorities, in the sense that teachers are encouraged to collaborate to develop a common understanding of the curriculum, the laws and regulations, and the assessment of student competence (Norwegian Directorate for Education and Training, 2016). As far as we are concerned, this aspect of a teacher's assessment competence is well worth emphasizing as a separate LAL component.

When it comes to the interrelationship between Taylor's dimensions, it was mentioned in the literature review above that *Principles and concepts* overlaps with the *Knowledge of theory* component. The reason for this is that concepts per se may be considered inherently theoretical (see e.g., Fulcher 2012). Thus, we suggest that "concepts" are removed from *Principles and concepts*. Instead, conceptual issues, such as validity and reliability, may be subsumed under *Knowledge of assessment theory*, as suggested above. Taylor's *Principles and concepts* dimension may then be relabelled *Principles* and reserved for knowledge of ethical issues and test fairness.

Furthermore, we find that *Scores and decision making* in Taylor's model also needs to be modified, as it partly overlaps with the *Language pedagogy* component. In the Literature review above, we pointed out that decision making not only relates to testing or summative decisions, but also to formative ones (see e.g, Bachman & Palmer, 2010, p. 197). Hence, it seems more pertinent to subsume "decision making" under *Language pedagogy*, as this component logically includes formative assessment. *Scores and decision making* can then be renamed *Scoring* and reserved for the awarding of grades, including the use of rating scales. Kremmel and Harding (2019) partly support this conclusion. They suggest that *Scores and decision making* should be renamed "Scoring and rating" and reserved for skills relating to scoring processes. They also propose that *Language pedagogy* be divided into two dimensions, "Assessment in language pedagogy" and "Washback and preparation". The former refers to feedback, diagnostic assessment, peer and self-assessment, as well as "how to use assessments to inform learning or teaching goals" (pp. 10, 16). The latter includes, for example, knowledge of "how assessments can influence teaching and learning in the classroom" (pp. 10, 17). This division seems pertinent, although it may not be quite clear what distinguishes feedback, peer and self-assessment, on the one hand, and how to use assessment to inform learning goals, on the other. When teachers use assessment information to adjust their teaching, cf. Wiliam's (2018) definition of formative assessment, above, that is something else than feedback, peer and self-assessment. We therefore believe that *Language pedagogy* may still be used as an overall category, as long as its content is sufficiently clarified. Table 3 gives an overview of what our suggested revision of Taylor's model implies.

TABLE 3
OVERVIEW OF TAYLOR'S MODEL AND A SUGGESTED REVISION OF IT.

Taylor's original dimensions	Suggestions for revised and new dimensions	Explanations for new and revised dimensions
Language pedagogy	*[No revision of title]*	Knowledge and skills related to formative assessment and teaching
Sociocultural values	*[No revision of title]*	[No revision of content]
Local practices	*[No revision of title]*	Knowledge of laws and regulations and local assessment criteria
Personal beliefs/attitudes	*[No revision of title]*	[No revision of content]
Scores and decision making	*Scoring*	Knowledge of scoring processes, including use of rating scales
Knowledge of theory	*Knowledge of assessment theory*	Knowledge of assessment theory, like for example validity theory
Technical skills	*[No revision of title]*	[No revision of content]
Principles and concepts	*Principles*	Knowledge of the ethics of testing and assessment and test fairness
	Disciplinary competence	Knowledge of language and language acquisition, curriculum knowledge and language skills
	Collaboration competence	Ability to contribute to creating a shared assessment culture with colleagues

As for the informants' answers to RQ2, the analysis shows that the teacher educators gave most of the items relating to Taylor's model high scores. This may be attributed to the propensity of respondents to deem everything important when presented with close-response items of this kind (Fulcher, 2012). However, the comparison of these responses with the answers to the open-ended questions revealed some general tendencies. Overall, the teacher educators agreed with Taylor that *Language pedagogy* (e.g., formative assessment) is an essential LAL component. Interestingly, this finding is not supported by Kremmel and Harding (2019), who did not find that "language pedagogy would be more important for this group than all other aspects of LAL" (p. 14). We speculate that this may be due to the high proportion of respondents from other contexts than the Norwegian one, where assessment for learning has had such influence in the past decade (Norwegian Directorate for Education and Training, 2015).

The teacher educators in the present study also agreed with Taylor that *Sociocultural values* (i.e. values in society), *Local practices* (e.g. curriculum), and *Personal beliefs/attitudes* (e.g. preconceptions), are important (cf., Fig. 1 and Fig. 2, above). However, contrary to Taylor, who judged *Scores and decision making* (e.g. grading), *Knowledge of theory* (e.g. language theory) and *Principles and concepts* (e.g. ethics) to be less important, our informants found these dimensions to be significant or very significant for teacher LAL (cf. Fig. 1 and Fig. 2). Finally, regarding *Technical skills* (e.g. statistics), the teacher educators on average thought that this dimension was less important than what Taylor hypothesized, although not trivial. This is supported by the findings in Kremmel and Harding (2019).

Combining our suggestion for a revised model with the responses the teachers gave to the questions in RQ1 and the items in RQ2, we end up with the following visualization of teacher LAL:
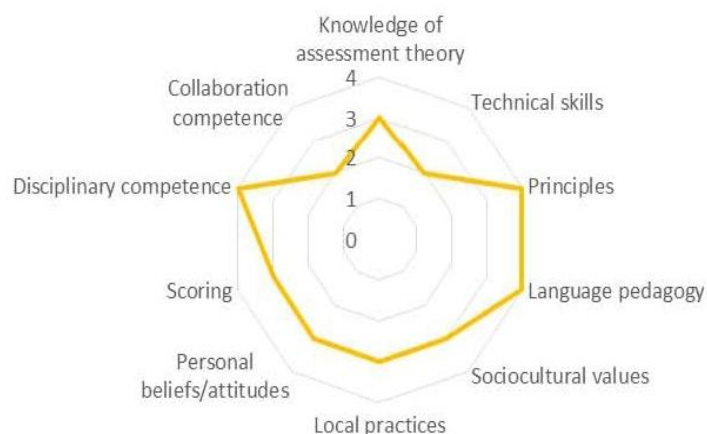

Figure 3: Revised version of Taylor's description including score results from the analysis

The values of the revised model are based on the responses to the items in RQ2, apart from *Collaboration competence*, which was not operationalized in the questionnaire, as it developed from the analysis of the open-ended responses to RQ1 post hoc. As only one informant mentioned this dimension, we have tentatively given it the value 2 in the description, i.e. fairly important. The explanations for each of the dimensions are given in Table 3, above.

A limitation of this study relates to the use of only one coder to analyse the informants' responses to the open-ended questions. The use of two or more coders would probably have increased the validity of the results. Another limitation involves the qualitative nature of the study, including its small number of informants, which obviously means that the results cannot be statistically generalized. However, the study's empirical findings have been supported elsewhere (e.g., Norwegian Directorate for Education and Training, 2019; Scarino, 2013; Xu & Brown, 2016). Also, the discussion of the theoretical characteristics of teacher LAL means that analytical generalizations from smaller samples may be warranted (Yin, 2016).

VII. CONCLUSIONS

The findings of this study have various implications for the development of teacher LAL. They point to the importance of strengthening teachers' disciplinary competence, especially regarding curriculum knowledge, as well as their ability to carry out good formative assessment practices. They also indicate that issues such as ethics in assessment and test fairness may need to be emphasized in teacher education and in-service programmes. In addition, they suggest that it is important that teachers develop skills in communicating and interacting with fellow colleagues in order to create a collaborative, professional community in which the focus is on establishing a shared assessment culture.

As for further avenues for research, we suggest that the perspectives of teacher educators and other stakeholders in other educational contexts are also studied. As the findings in this investigation have supported Xu and Brown's (2016) claim that LAL is contextual rather than universal, it would be interesting to get to know what language teaching experts in other countries have to say about the relevance of Taylor's, or Kremmel and Harding's (2019), models. Similarly, it would be intriguing to bring in the perspectives of teachers, both in Norway and elsewhere, to see whether their views align with the viewpoints of the teacher educators in this study.

REFERENCES

[1] Abell, S. K., & Siegel, M. A. (2011). Assessment literacy: What science teachers need to know and be able to do. In D. Corrigan, J. Dillon, & R. Gunstone (Eds.), *The professional knowledge base of science teaching*. Dordrecht: Springer.

[2] Bachman, L., & Palmer, A. (2010). Language assessment in practice. Oxford: Oxford University Press.

[3] Baker, B. A. (2016). Language Assessment Literacy as Professional Competence: The Case of Canadian Admissions Decision Makers. *The Canadian Journal of Applied Linguistics, 19*(1), 63-83.

[4] Baker, B. A., & Riches, C. (2018). The development of EFL examinations in Haiti: Collaboration and language assessment literacy development. *Language Testing, 35*(4), 557-581. doi:10.1177/0265532217716732.

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

[6] Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability(formerly: Journal of Personnel Evaluation in Education), 21*(1), 5-31. doi:10.1007/s11092-008-9068-5.

[7] Bloom, B. S. (1969). Some theoretical issues relating to educational evaluation. In H. G. Richey & R. W. Tyler (Eds.), *Educational evaluation: New roles, new means, part 2* (Vol. 68, pp. 26-50). Chicago: University of Chicago Press.

[8] Brindley, G. (2001). Language assessment and professional development. In C. Elder, A. Brown, K. Hill, N. Iwashita, T. Lumley, T. McNamara, & K. O'Loughlin (Eds.), *Experimenting with uncertainty: Essays in honour of Alan Davies* (pp. 126-136). Cambridge University: Cambridge University Press.

[9] Bryman, A. (2016). Social research methods (5th ed.). Oxford: Oxford University Press.

[10] Coombs, A. J. (2017). Teacher educators' approaches to assessment. (Master's thesis). Queen's University, Ontario, Canada.

[11] Creswell, J. W. (2013). Qualitative inquiry & research design: Choosing among five approaches. Los Angeles: Sage.

[12] Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing, 25*(3), 327-347. doi:10.1177/0265532208090156.

[13] Deygers, B., & Malone, M. E. (2019). Language assessment literacy in university admission policies, or the dialogue that isn't. *Language Testing, 36*(3), 347-368. doi:10.1177/0265532219826390.

[14] Fulcher, G. (2012). Assessment Literacy for the Language Classroom. *Language Assessment Quarterly, 9*(2), 113-132. doi:10.1080/15434303.2011.642041.

[15] Fulcher, G., & Davidson, F. (2007). Language testing and assessment: An advanced resource book. Oxon: Routledge.

[16] Giraldo, F. (2018). Language assessment literacy: Implications for language teachers. *Profile: Issues in Teachers' Professional Development, 20*(1), 179-195. doi: https://doi.org/10.15446/profile.v20n1.62089.

[17] Harding, L., & Kremmel, B. (2016). Teacher assessment literacy and professional development. In D. Tsagari & J. Banerjee (Eds.), *Handbook of Second Language Assessment* (pp. 413-427). De Gruyter Mouton.

[18] Hildén, R., & Fröjdendahl, B. (2018). The dawn of assessment literacy – exploring the conceptions of Finnish student teachers in foreign languages. *Apples: Journal of Applied Language Studies, 12*(1), 1-24. doi:10.17011/apples/urn.201802201542.

[19] IBM Support. (no date). Transforming different Likert scales to a common scale. https://www.ibm.com/support/pages/node/422073 (accessed 29/1/2020).

[20] Inbar-Lourie, O. (2008). Language assessment culture. In E. Shohamy & N. H. Hornberger (Eds.), *Encyclopedia of Language and Education* (2nd ed., Vol. 7, pp. 285-289). New York: Springer.

[21] Inbar-Lourie, O. (2013). Guest Editorial to the special issue on language assessment literacy. *Language Testing, 30*(3), 301-307. doi:10.1177/0265532213480126.

[22] Inbar-Lourie, O. (2017). Language assessment literacy. In E. Shohamy, I. G. Or, & S. May (Eds.), *Language Testing and Assessment* (3rd ed.). Cham: Springer.

[23] Jamieson, S. (2004). Likert scales: how to (ab)use them. *Medical education, 28*, 1212-1218. Retrieved from file:///C:/Users/hbo/Downloads/1LikertScales.pdf

[24] JCSEE (Joint Committee on Standards for Educational Evaluation). (2015). The Classroom Assessment Standards for PreK-12 Teachers. https://www.amazon.com/Classroom-Assessment-Standards-PreK-12-Teachers-ebook/dp/B00V6C9RVO (accessed 14/2/2020).

[25] Jeong, H. (2013). Defining assessment literacy: Is it different for language testers and non-language testers? *Language Testing, 30*(3), 345-362. doi:10.1177/0265532213480334.

[26] Jin, Y. (2010). The place of language testing and assessment in the professional preparation of foreign language teachers in China. *Language Testing, 27*(4), 555-584. doi:10.1177/0265532209351431.

[27] Kremmel, B., & Harding, L. (2019). Towards a Comprehensive, Empirical Model of Language Assessment Literacy across Stakeholder Groups: Developing the Language Assessment Literacy Survey. *Language Assessment Quarterly*, 1-21. doi:10.1080/15434303.2019.1674855.

[28] Lam, R. (2014). Language assessment training in Hong Kong: Implications for language assessment literacy. *Language Testing, 32*(2), 169-197. doi:10.1177/0265532214554321.

[29] López Mendoza, A. A., & Bernal Arandia, R. (2009). Language Testing in Colombia: A Call for More Teacher Education and Teacher Training in Language Assessment. *Profile Issues in Teachers` Professional Development, 11*, 55-70. Retrieved from http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S1657-07902009000200005&nrm=iso.

[30] Malone, M. E. (2013). The essentials of assessment literacy: Contrasts between testers and users. *Language Testing, 30*(3), 329-344. doi:10.1177/0265532213480129.

[31] McNamara, T., & Roever, C. (2006). Language Testing: The social dimension. London: Blackwell.

[32] Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). Qualitative data analysis: A methods sourcebook (3rd ed.). Thousand Oaks: Sage.

[33] Norwegian Directorate for Education and Training. (2015). Vurdering for læring - om satsingen [Assessment for learning - a note on the priority area]. https://www.udir.no/laring-og-trivsel/vurdering/nasjonal-satsing/om-satsingen/ (accessed 15/1/2020).

[34] Norwegian Directorate for Education and Training. (2016). Samarbeid og tolkningsfellesskap [Collaboration and a shared assessment culture]. https://www.udir.no/laring-og-trivsel/vurdering/om-vurdering/samarbeid-om-vurdering/ (accessed 13/1/2020).

[35] Norwegian Directorate for Education and Training. (2017). Overordnet del - Verdier og prinsipper i grunnopplæringen [Core curriculum - values and principles in primary and lower secondary education]. https://www.udir.no/lk20/overordnet-del?kode=eng01-04&lang=nob (accessed 24/1/2020).

[36] Norwegian Directorate for Education and Training. (2019). Erfaringer fra nasjonal satsing på vurdering for læring (2010-2018) [Experiences from the national priority programme assessment for learning (2018-2019). https://www.udir.no/tall-og-forskning/finn-forskning/rapporter/erfaringer-fra-nasjonal-satsing-pa-vurdering-for-laring-2010-2018/ (accessed 4/1/2020).

[37] Norwegian Ministry of Education and Research. (2006/2019). Forskrift til Opplæringslova [Regulations to the Education Act]. https://lovdata.no/dokument/SF/forskrift/2006-06-23-724 (accessed 24/1/2020).

[38] O'Loughlin, K. (2013). Developing the assessment literacy of university proficiency test users. *Language Testing, 30*(3), 363-380. doi:10.1177/0265532213480336.

[39] OED. (2002). Model. In *Oxford English Dictionary* (3rd ed.). Oxford: Oxford University Press.

[40] Pill, J., & Harding, L. (2013). Defining the language assessment literacy gap: Evidence from a parliamentary inquiry. *Language Testing, 30*(3), 381-402. doi:10.1177/0265532213480337.

[41] Popham, W. J. (2011). Assessment literacy overlooked: A teacher educator's confession. *Teacher Educator, 46*, 265–273. doi:10.1080/08878730.2011.605048

[42] Scarino, A. (2013). Language assessment literacy as self-awareness: Understanding the role of interpretation in assessment and in teacher learning. *Language Testing, 30*(3), 309-327. doi:10.1177/0265532213480128.

[43] Shulman, L. (1987). Knowledge and Teaching: Foundations of the New Reform. *Harvard Educational Review, 57*(1), 1-23. doi:10.17763/haer.57.1.j463w79r56455411.

[44] Stiggins, R. J. (1991). Assessment Literacy. *The Phi Delta Kappan, 72*(7), 534-539. Retrieved from http://www.jstor.org/stable/20404455.

[45] Taylor, L. (2009). Developing assessment literacy. *Annual Review of Applied Linguistics, 29*, 21-36. Retrieved from file:///C:/Users/hbo/Downloads/DEVELOPING_ASSESSMENT_LITERACY.pdf.

[46] Taylor, L. (2013). Communicating the theory, practice and principles of language testing to test stakeholders: Some reflections. *Language Testing, 30*(3), 403-412. doi:10.1177/0265532213480338.

[47] Villa Larenas, S. (2017). Language assessment literacy of EFL teacher trainers. Paper presented at the 39th Language Testing Research Colloquium, Universidad de los Andes, Bogotá Colombia.

[48] Vogt, K., & Tsagari, D. (2014). Assessment Literacy of Foreign Language Teachers: Findings of a European Study. *Language Assessment Quarterly, 11*(4), 374-402. doi:10.1080/15434303.2014.960046.

[49] Wiliam, D. (2018). Embedded formative assessment (2nd ed.). Bloomington: Solution Tree Press.

[50] Xu, Y., & Brown, G. T. L. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education, 58*, 149-162. doi:https://doi.org/10.1016/j.tate.2016.05.010.

[51] Yan, X., Zhang, C., & Fan, J. J. (2018). "Assessment knowledge is important, but …": How contextual and experiential factors mediate assessment practice and training needs of language teachers. *System, 74*, 158-168. doi:https://doi.org/10.1016/j.system.2018.03.003.

[52] Yin, R. K. (2016). Qualitative research from start to finish (2nd ed.). New York: The Guilford Press.

Henrik Bøhn, PhD, is Associate Professor at the Department of Languages and Literature, University of South-Eastern Norway, Norway. He worked as an English teacher in upper secondary school in Norway for three years before becoming a lecturer of English at the university level in 2000. Since 2008 he has mainly worked as an English language teacher educator. His research interests include language assessment, language acquisition and intercultural communication.

Dina Tsagari, PhD, is Professor at the Department of Primary and Secondary Teacher Education, Oslo Metropolitan University, Norway. She worked for the University of Cyprus, Greek Open University and Polytechnic University of Hong Kong. Her research interests include language testing and assessment, materials design and evaluation, differentiated instruction, multilingualism, distance education and learning difficulties. She is the editor and author of volumes, journal papers, and book chapters. She coordinates the CBLA SIG – EALTA, EnA OsloMet and is involved in EU-funded and other research projects (e.g. DINGLE, ENRICH, TALE, DysTEFL, PALM)