MDPI

Article

The Impact of Student Teachers' Pre-Existing Conceptions of Assessment on the Development of Language Assessment Literacy within an LTA Course

Olga Kvasova

Department of Methods of Teaching Ukrainian and Foreign Languages and Literature, Taras Shevchenko, National University of Kyiv, 016016 Kyiv, Ukraine; ualta@ukr.net

Abstract: The paper proposes the preparation of a new generation of assessment literate teachers. The issues of student assessment literacy and, more specifically, prospective language teacher assessment literacy have not been sufficiently investigated as of yet, although research into the topic seems to have gained momentum. Recent studies state that the assessment literacy of teachers is essentially affected by their pre-existing conceptions of assessments), and teacher education should integrate shaping such conceptions into courses; the process of shaping conceptions is quite long and, because it is time-consuming, it may deter assessment literacy building. The current study explores the conceptions of the assessments shaped by prospective teachers within a general English course. The two major conceptions of assessment, relevant for the framework of teaching general English to second-year student teachers of English, are the understanding of feedback and knowledge of assessment construct and criteria. The findings of the study in this cohort of students of the particular course in language assessment shows that the students' progress was considerably higher than that of a comparison group in the previous 2020 study. The author suggests two types of AL, i.e., student and prospective teacher assessment literacy.

Keywords: student assessment literacy; conceptions of assessment; feedback and knowledge of assessment criteria; language assessment literacy of prospective teachers

1. Introduction

Language assessment literacy (LAL) has moved to the centre stage in the field of language testing and assessment (LTA). A number of prominent theoretical studies, addressing LAL's conceptual frameworks and contextualised conceptions, have marked the past decade (Fulcher 2012; Inbar-Lourie 2013; Taylor 2013; Pill and Harding 2013; Vogt and Tsagari 2014; Coombe et al. 2020; Bøhn and Tsagari 2021). As a result, courses in LTA have increased and gained higher status within pre-service teacher training curricula (Fröjdendahl 2018; Lam 2015; Odo 2016; Solnyshkina et al. 2016; Ukrayinska 2018; Kvasova 2020; Martínez Marín and Vélez 2021).

The assessment literacy of students, in general, as well as that of prospective teachers, in particular, is all the more addressed in literature, thus responding to the emerging needs in society (Smith et al. 2013; Massey et al. 2020; Noble 2021). The trend is explained by an increasing interest in equipping novice teachers with expertise in performing assessments as part of their job. The influx of assessment-literate teachers in the profession will allow for the redirection of administrative resources from in-service training in the *stable knowledge base* acquisition (Brown and Bailey 2008) to more demanding topics of assessment, such as online classroom assessment, alternative assessment, etc.

Unlike practicing teachers, pre-service trainees in LTA do not have hands-on experience in teaching/assessment to reflect and rely on. As Kvasova (2020) states, undergraduates do not approach LTA from a teacher perspective, but associate themselves primarily

with assesses, drawing mostly on their actual experience as students being assessed. This experience, together with beliefs, values, and attitudes, shapes student teachers' pre-existing conceptions of assessment, which can be positive/negative and dynamic/persistent. In any case, the conceptions take a much longer time to develop than the duration of an academic course in LTA (Deneen and Brown 2016). The development process is affected by multiple factors, such as practicum (Xu and He 2019), context, and career stage (Coombe et al. 2020), as well as personal development and enhanced educational and professional levels. We may assume that the popular saying 'teachers teach as they were taught' may relate to the pre-existing conceptions in novice teachers' consciousness, which guide them in their practice, despite the newly received formal training. As Smith et al. (2014) found, the teachers' assessment beliefs were formed by their past experiences, rather than by the content they were taught.

Although cognitive psychology and education research widely recognize the influence of pre-existing knowledge and experiences in shaping the cognition and behavior of prospective instructors (e.g., see Oleson and Hora 2013), the above quoted claim has not been supported empirically as of yet. In fact, Cox provided evidence that "there is a significant difference between how teachers teach and how they were taught during their own educational experience" (Cox 2014, p. ii), which needs further investigation. Regarding LAL, Grainger and Adie argue that "learning to be an assessor through being assessed can be problematic, particularly when working between the different levels of education" (Grainger and Adie 2014, p. 90). This seems to apply mostly to the faculty, who have not received formal training in pedagogical methods (Oleson and Hora 2013) and, therefore, are unable to overcome the intuitive and imitative 'folkways of teaching' (Lortie 1975).

The current study focuses on the impact of student teachers' pre-existing conceptions of assessment on the more conscious formal training in LAL. More particularly, the pre-existing conceptions of assessment developed tacitly in a General English (GE) course classroom are examined and how these influence LAL development within a course in LTA.

2. Literature Review

Assessment literacy consists of knowledge and attitudes, or conceptions (Deneen and Brown 2011), similar to any other competence. Following Deneen and Brown (2016), the term conception is, in this paper, inclusive of attitudes, perceptions, dispositions, and other terms that suggest belief about a phenomenon. Conceptions of assessment (CoA) are key elements of assessment competence (Hill and Eyers 2016) and constitute an important part of a teacher's assessment identity (Looney et al. 2018). They guide teacher assessment practices (Barnes et al. 2014; Brown 2008) and play a pivotal role in achieving assessment literacy (AL) (Deneen and Brown 2016).

Deneen and Brown (2016) argue that CoA have significant implications for teacher education. Their research shows that the theoretical knowledge acquired, and the practical assessment skills developed are not sufficient for the overall development of AL. The authors reiterate that shaping conceptions requires a lot of time and students' gains in knowledge and skills across a brief training course may not, regrettably, be accompanied by enhanced CoA, which eventually impedes students' AL achievement.

Smith et al. (2013) define student AL as "students' understanding of the rules surrounding assessment in their course context, their use of assessment tasks to monitor, or further, their learning, and their ability to work with the guidelines on standards in their context to produce work of a predictable standard" (p. 46). As Noble (2021) recaps this definition, AL is "students' ability to understand the purpose and processes of assessment, and accurately judge their own work" (p. 1). It presupposes (1) understanding key assessment terminology and different assessment methods and procedures, in the context of university, (2) being familiar with performance standards and criteria used to judge student work, and

(3) being able to judge or evaluate own/peers' performances on assessment tasks marked by criteria and standards (Noble 2021).

In a similar vein to Smith et al. (2013) treatment of student AL, O'Donovan et al. (2004) focus on the development of student understanding of assessment standards and approaches to assessment knowledge transfer. This development may proceed in either an explicit or tacit way, with varied student engagement—passive or active. The authors claim that the future lies in the active student engagement, facilitated by the social constructivist approach (through active student engagement in formal processes devised to communicate tacit knowledge of standards) or the 'Cultivated' community of practice approach (wherein tacit standards communicated through participation in informal knowledge exchange networks 'are seeded' by specific activities). The authors advocate the 'social constructivist process model' of assessment, where "students are actively engaged with every stage of the assessment process in order that they truly understand the complex and contextual requirements of assessment praxis, particularly the criteria and standards being applied, and how they may subsequently produce better work" (O'Donovan et al. 2004, p. 208).

The scholars argue that developing tacit dimensions of knowledge is more effective than the transfer of explicit information, since learners obviously need to be engaged in meaningful activities to be able to construct meaning for themselves. O'Donovan et al. (2008) contend that tacit knowledge is built up through socialisation and practice and can be internalized by the student to shape their understanding of standards and expectations. The implication from the above is that, in order to come to grips with assessment standards, students need not only to gain the knowledge transferred explicitly, but also to take advantage of tacit acquisition of knowledge, whenever this is possible, including the situations created by teachers and embedded in the instruction.

Finally, O'Donovan et al. propose facilitating learners' 'pedagogical intelligence' by inducting them, not only into the discourse of their chosen discipline, but also into the discourse of learning and teaching. The authors concur with Hutchings (2005), who proposes that "pedagogical intelligence would require students not only to be able to critically evaluate their own learning, but also their learning environment including the teaching" (quoted in O'Donovan et al. 2008, p. 214). The idea to engage students, especially student teachers, with concepts, principles, and technicalities of assessment in a tacit paradigm, within any discipline of their study, appears quite constructive. It resonates with Tenet 5 (integrating assessment literacy into course design), as proposed by the Higher Education Academy Project Report (2012). The tenet mentions that "Active engagement with assessment standards needs to be an integral and seamless part of course design and the learning process in order to allow students to develop their own, internalized conceptions of standards and to monitor and supervise their own learning" (p. 21).

Given that significant time and effort are required for serious conceptual shift to happen, it may seem quite logical for student teacher assessment conceptions to start shaping prior to the beginning of student teachers' formal training in LTA. There is evidence that teachers' CoA arise from their experiences of being assessed as learners (Hill and Eyers 2016) and prior personal experiences of being assessed before and during teacher education play a significant role in structuring their assessment conceptions (Crossman 2007; Smith et al. 2014). It is worth mentioning the experiences of apprenticeship of observation (Lortie 1975) that involve subconscious learning about teaching through observing their own teachers' teaching. These significantly impact pre-service teachers' initial conceptions and account for their teaching practice, as they offer insights into what to do as teachers (Knapp 2012; Boyd et al. 2013; quoted in Xu and He 2019).

Emphasizing the role of CoA within AL, Deneen and Brown argue that "courses in assessment, perhaps more than any other aspect of teacher education, must address preexisting conceptions and beliefs and their causes" (Deneen and Brown 2016, p. 14). Following this line of thought, the authors call for a sustained program-level engagement with CoA that will lead to innovating the curricula content and structure and modernizing the concept of AL itself.

As far as LAL is concerned, after being identified by Davies (2008) as knowledge, skills, and principles that stakeholders are required to master, in order to perform assessment tasks, the concept has evolved significantly. More extended definitions of LAL are proposed by Fulcher (2012), Inbar-Lourie (2013), Vogt and Tsagari (2014), Giraldo (2018), and others. One of the most recent generic definitions of LAL is offered by Coombe et al. (2020), who, referring to Inbar-Lourie (2008), Pill and Harding (2013), and Stiggins (1999), interpreted LAL as "a repertoire of competences, knowledge of using assessment methods, and applying suitable tools in an appropriate time that enables an individual to understand, assess, construct language tests, and analyze test data" (p. 2).

While the "knowledge-skills-principles" paradigm is frequently referred to in the definitions of the above-mentioned and other authors, Scarino (2013) contends that there is one more integral part of teacher LAL—their self-awareness as assessors. Drawing on the impact of sociocultural theories of learning on assessment practices in second language education, she argues for the need to enable teachers to explore and evaluate their own preconceptions, beliefs, personal theories, and experience, so as to enhance critical awareness of their own assessment practices and students' second language learning.

The assessor's personal beliefs and attitudes occupy a specific area within differential profiles of LAL, as proposed by Taylor (2013). In the profile for professional language testers, these aspects are in an equally high place with all other aspects, such as knowledge, principles, etc. In the profile for language teachers, personal beliefs/attitudes are ranked second, together with technical skills, sociocultural values, and local practices, following the top-ranked language pedagogy. In the second position, there are also the personal beliefs/attitudes in the version of Taylor's description, revised by Bøhn and Tsagari (2021), based on their empirical study of teacher educator conceptions of LAL. Does the inclusion of personal beliefs/attitudes in LAL profiles mean that they should be included in the LAL construct and taken into thorough consideration? In any case, we cannot deny that teachers' contexts for language assessment should contribute to the meaning of LAL, given that one's own preconceptions, understandings, and opinions may "inform [one's] conceptualizations, interpretations, judgments and decisions in assessment" (Scarino 2013, p. 109).

Before voices were raised for the necessity of building LAL in various stakeholders (Fulcher 2012; Pill and Harding 2013; Taylor 2013; Kremmel and Harding 2019), LAL was mostly associated with the LAL of two major stakeholder groups—language testers and teachers (Volante and Fazio 2007; Brookhart 2011; Fulcher 2012; Vogt and Tsagari 2014; DeLuca et al. 2016; Xu and Brown 2016). According to Fulcher (2012), assessment literacy for language testers comprises the knowledge, skills, and abilities required to design, develop, maintain, or evaluate large-scale standardized and/or classroom-based tests, familiarity with test processes, and awareness of principles and concepts that guide and underpin practice, including ethics and codes of practice. As for teacher assessment literacy, it has been considered mostly from the perspective of in-service teacher training in LTA, based on the needs that had been empirically explored (Hasselgreen et al. 2004; Fulcher 2012; Vogt and Tsagari 2014; Kvasova and Kavytska 2014; Massey et al. 2020; Bøhn and Tsagari 2021). Teachers' beliefs and perceptions concerning the relevance of language assessment and the challenges it poses have also been studied (Wach 2012; Tsagari and Vogt 2017; Giraldo 2019; Vogt et al. 2020). Moreover, some works report on planning and implementing in-service teacher training courses in LTA (e.g., Tsagari et al. 2018).

Malone (2017) claims that it is time to expand the understanding of LAL by including students within the stakeholders and that, so far, "only limited research . . . has explored students' understanding of assessment, assessment results and their own performances" (p. 9). In fact, Malone (2017) warns against the erroneous transfer of practicing teachers' understanding of major assessment concepts onto students during training, since a lack of relevant research reflects that "language assessment literacy efforts for students represent still more complexity for the field" (p. 83). The past few years, research into student assessment literacy seems to be gaining momentum through studies that discuss the organization of training courses for prospective/pre-service teachers (Beziat and Coleman

2015; Lam 2015; Yetkin 2018; Ukrayinska 2018; Hildén and Fröjdendahl 2018; Fröjdendahl 2018; Giraldo and Murcia 2019; Kvasova 2020). The common element in these studies is that their authors proceed from identifying the gaps in in-service teacher LAL, in order to define the training needs aimed at bridging these gaps and, finally, building up curricula for pre-service training courses. Proceeding, in this case, from teacher training needs, as a starting point, is viewed as absolutely justified. It is also true, however, that the differences between the objectives of courses aimed at two cohorts of trainees are either ignored or minimized; therefore, the curriculum for practicing teachers is considered easily applicable to prospective teachers' training (e.g., Anam and Putri 2021).

Tsagari (2017) argues that the acquisition and implementation of LAL is context-specific, depending on social and educational values, as well as the beliefs shared within a particular assessment culture. Besides, LAL is not static, but dynamic, since it is affected by the teachers' individual perceptions and knowledge about learning and assessment. Acknowledging the importance of contextualization of LAL, we agree that LAL of two major stakeholder groups—teachers and students—is specific. It is situated in particular contexts and determined by teachers and/or students' social roles, education level, status, age, maturity, etc. Thus, specificity of LAL presupposes "identification of assessment priorities and the development of assessment training strategies that are contextually situated within effective modes of training" (Tsagari 2017, p. 83).

After looking into research in this field, what seems to be lacking is a comprehensive definition of student teacher LAL. This paper attempts to contribute in this respect.

3. Current Research
3.1. Defining Students' Pre-Existing Conceptions of Assessment

According to Brown (2008), students conceive assessment in four ways, i.e., the assessment makes students accountable, is irrelevant because it is bad or unfair, improves the quality of learning, and is enjoyable. The methodology to research the CoA has been employed multiple times by researchers (Brown and Remesal 2012; Brown et al. 2014; Opre 2015; Yetkin 2018; Lutovac and Flores 2021; Herrera Mosquera and Meléndez 2021). However, the purpose of contextualizing student CoA, pursued in this study, calls for a finer adjustment to a particular teaching/learning environment (the course curriculum, characteristics of the cohort of students, pedagogical professionalization), and a specification of CoA relevant for this particular context. The latter should be fundamental and transferable from current learning to prospective teaching practice.

It has become clear, from research so far, that the central idea of assessment itself and LAL is that of feedback. Feedback is where the fusion of learning and assessment occurs or the learner and assessor most closely interact during instruction. As learners, students conceive of assessment feedback as major evidence of progress in their learning. As prospective teachers, they need to recognize what exactly makes each form of feedback meaningful and be aware of what ways feedback can be beneficial to learners for promoting learning. Overall, good feedback should be prospective (i.e., guide and encourage learners to improve learning), specific (i.e., indicate which aspects of performance need improvement), timely (i.e., offered promptly with opportunities for learners to improve before the unit ends), and directly related to learning goals (i.e., based on curriculum and ensuring that learners understand the criteria) (Green 2014). Based on the above-mentioned, as well as other premises (Evans 2013; Tsagari et al. 2018), feedback is understood in this study as:

- Promoting better learning;
- Helping to identify positive and negative features of performance;
- Contributing to direct further learning at particular language aspects;
- Promoting self-assessment.

To benefit from feedback, learners and/or prospective teachers need to know, understand, and be able "to interpret the criteria [of assessment] and use them to assess the work of other learners and, eventually, to assess their own work" (Green 2014, p. 93). Therefore, the assessment should be transparent to learners with the criteria utilized, consciously,

though naturally, and transparent, helping learners to understand why one's work meets (or does not meet) the criteria for good performance and how it can be improved. If explained properly, the criteria should direct learners towards more conscious and controlled performance and enable subtler feedback interpretation, thus promoting more targeted remedial work and/or further performance. In the light of the above speculations, it appears well-reasoned for this study to define conception as "conscious knowledge of assessment construct and criteria" and specify it as the knowledge of:

- What skills are assessed in each task;
- What task types each test consists of;
- What criteria are applied in assessment of productive skills;
- What exactly each criterion measures.

Given that student AL is viewed as understanding the purposes, procedures, and ability to judge their own work, it is assumed in this study that the types of pre-existing student CoA relevant for the specific context of the GE course running in Year Two are (a) the understanding of feedback and (b) the conscious knowledge of assessment construct and criteria.

As was argued above, pre-existing CoA are shaped during students' prior learning experiences. In this study, the students' experiences gained in the GE course, that immediately precedes the beginning of formal study of LTA, are examined and considered as the prior learning experience that generated the understanding of feedback and knowledge of construct and criteria. Given the focus on the timespan during which particular activities envisaged by the GE curriculum (year, term, course book) were performed, the state of CoA is viewed as the contextualized function of the GE course implementation (see detailed description of the course in Section 4.1).

3.2. Research Framework

The framework of the current study encompasses two academic courses: GE and LTA. Both courses were conducted in two consecutive terms (Term 1 and Term 2), in the same academic year and by the same instructor—the author of this article. She has extensive experience in teaching English at the university level and has been actively engaged in research into LTA (e.g., Kvasova 2009; Kvasova and Kavytska 2014; Kvasova 2020). Nevertheless, the study needs to be viewed as teacher action research, with certain limitations imposed by the inevitable bias from the dual role of the researcher.

3.2.1. Participants

There are two sets of student participants in the study. The first set consisted of 18 2nd-year students in Term 1 and the same 18 students, plus a new one (in total 19), in Term 2. The second set was the comparison group, consisting of 21 students, who completed their second year several years before the current study took place. They participated in the authors' 2020 research, whose results are used as a benchmark in the current study.

Each of the sets include all students of Year Two enrolled in the programme in the respective years. The variable between the two sets was that the cohort of Set 2 had not been previously taught the GE course by the teacher–researcher; therefore, the development of pre-existing CoA within the GE course had not been examined.

Both sets (40 students in total) were females, aged 18, and native speakers of Ukrainian Their major at university was in teaching L1, with the additional specialization in TEFL. The educational programme that the participants pursue lasts four years, preparing them for teaching Ukrainian and literature, as well as a foreign language in secondary school. The curriculum includes a range of pedagogical courses, such as general pedagogy, psychology, methods of teaching L1 and L2, and literatures, thus providing a solid theoretical foundation for LAL acquisition within the LTA course.

3.2.2. Materials

The GE course is based on "A Way to Success", a multi-skill course book for first year university students who major in English (Tuchyna et al. 2019). It takes learners from CEFR level B 1+, which is the entry level to study at the university in Ukraine, presumably to B1.1. Since the students participating in the study do not major in English, the course book level appears appropriate for their second year of study. The number of contact hours for Term 1 is 56 and for Term 2 the number was 63. The course ends with an exam, upon completion of Term 2.

The course in LTA is divided in two parts: half of it is dedicated to L1 assessment and delivered in L1, and the other half focuses on LTA in teaching foreign languages and is delivered in English. The LTA course runs in Term 2; its part, delivered in English, spans 14 contact hours and ends with a summative test. The theoretical part of the course is based on the coursebook written in L1 (Kvasova 2009) and multiple articles on relevant issues published in the national ELT journals.

3.2.3. Research Purpose and Questions

The idea for this research was conceived when the teacher–researcher taught both courses, with the same cohort of students within two consecutive terms in the academic year 2019–2020. It seemed interesting to look into the possible impact of teaching by an assessment-minded instructor on the efficiency of students' further formal training in LTA.

According to the research design, the GE course was the site for building the students' CoA in a tacit way, similar to the 'apprenticeship of observation', though with some specific emphasis on implementing the assessments regularly placed by the instructor. The objective of the LTA course was to build student LAL. The study aims to resolve two research questions:

RQ1. Did the trainees develop such CoA as "understanding of feedback" and "knowledge of assessment construct and criteria" during the GE course that precedes the LTA course?

RQ2. Did the CoA developed within the GE course impact the students' LAL developed within the LTA course?

The purpose of the study, therefore, is to explore the possibility and conditions of CoA tacit development within a language course, when the assessment is properly integrated in the teaching/learning process. It is significant for teacher education, both in terms of improving the quality of language learning and learning in the profession-oriented course in LTA.

3.2.4. Methods

The methodology followed in this study includes two questionnaires, namely Questionnaires 1 and 2, as well as an exit test. Questionnaire 1 was administered in Term 1, upon completion of the GE course; Questionnaire 2, along with the exit test, was administered in Term 2, upon completion of the LTA course.

Questionnaire 1 was developed by the teacher–researcher to investigate the state of the two CoA before the formal training in LTA began. It was piloted with the help of two colleagues who were teaching English to the same cohort of learners and well-familiar with their aptitude.

Q.1–Q.10 are Likert-type, graded, four-option questions. Q.1–5 are focused on the conception of feedback, whereas Q.6–10 are based on the conception of knowledge of criteria. The questions are formulated based on the five activities that were performed and assessed, according to the GE course curriculum (individual long turn, role plays, revision translation, writing tasks, and tests).

Questionnaire 2 consists of six Likert-type questions with five traditional options. The objective of this questionnaire is to elicit if the respondents perceived any impact of taking the GE course on the learning outcomes of the LTA course, as well as whether this impact was positive. To compile the questionnaire was rather an ethical challenge for

the researcher who was also the teacher in both courses. Therefore, she tried to evade formulating questions in a direct manner, focusing on the same issues several times from different perspectives and preferably indirectly. (See Questionnaire 1 translated into English in Appendix A).

Questionnaire 2 was piloted first with the help of two fellow-teachers and then administered at the end of Term 2 after the course in LTA was delivered. It was administered via Google forms.

Both questionnaires were offered in L1 to ensure its clarity to everyone irrespective of English language proficiency.

The exit test consists of 25 selected response items, focused on knowledge of the major principles of classroom test development: Task 1, test development cycle; Task 2, aspects of test usefulness; Task 3, testing techniques and rubrics. The test was offered in English (See Appendix B). It had been developed by the author of the paper and pre-tested with the help of three colleagues before the 2020 study. The test is used by the instructor as summative assessment at the end of the LTA course each year. In 2020 and 2021, the test was administered online.

4. Data Collection and Analysis

4.1. Exploring GE Course Delivery

The curriculum of the multi-skill GE course for Term 1, which comprises two modules, addresses mainly general interest topics, such as student life (studies and pastime), appearance, and personal characteristics. The teacher's book, similar to most teacher's books, does not contain any assessment materials, e.g., quizzes, tests, and rating scales, to assess speaking and writing. Course activities include a variety of classroom activities such as, questioning, written home assignment, whole-group discussion, paired/small team discussion, role plays, quizzes, revision translation, individual long turn, written tests, and writing tasks. The assessment phase of the activities is planned by the teacher, with respect to the objectives of the activities and their formats. Therefore, revision translation, individual long turn (oral test), written test, and a writing task are viewed as summative assessments, whereas the rest of the activities are assessed continuously and formatively. Thus, across the term, the students are exposed to a variety of formative and summative assessment techniques, with the exception of peer- and self-assessment, which, according to the curriculum, are offered in Term 2. The course is completed with administering Questionnaire 1.

Since the instructor's research and professional interests lay within the field of LTA, she conducted assessments 'by the assessment book', drawing from the works of Coombe et al. (2007), Green (2014), and Tsagari et al. (2018). In particular, she made a point of explaining the tasks to be performed/assessed in a meticulous fashion, indicating such necessary parameters as timing, score, example, answer grid (for the test format of tasks), and the assessment criteria (for productive skills). She would also strive for the test papers to be appropriately laid-out and have positive face validity, thus seeking to develop the learners' diligence while using and explaining the tests. Regarding the assessment of speaking and writing, she made a point of introducing criteria, and even rating scales, to students and used and explained the necessary metalanguage, e.g., 'coherence and cohesion'. Students' performances were commented on throughout, with comments frequently containing suggestions as to what and how it should be improved.

As the classes during the academic year were conducted mostly online, due to the COVID-19 pandemic, the teacher made it a rule to assess students' written assignments via Google classroom, providing corrections and suggestions wherever possible. When assessing written compositions electronically, the teacher would indicate relevant grids in the rating scale to make the scores—per criterion and total—more explicit and comprehensive. Additionally, the final score for written works was normally delayed, providing the students with an option to edit their work, thus making the most of the teacher's written feedback.

It is worth noting that, although quite concentrated, the efforts the teacher put into instilling the norms of conducting assessments properly were far from being an intervention, as they did not stand out purposely and disrupt the natural flow of classroom activities. This was confirmed in conversations with the instructor, who was teaching the same curriculum in the other group of second year students. Additionally, the teaching and assessment process was monitored by the education managers, who did not notice any deviations from the normal instructional procedure. However, there was another purpose of conducting assessment by the book, beyond ensuring fair assessment within the course. The teacher meant for the students to observe assessment procedures repeatedly, systemically, and in compliance with testing principles or 'cornerstones' (termed so by Coombe et al. 2007). As was hypothesized, the students could develop certain preconceptions of good assessment practice in a tacit way that would facilitate their further formal acquisition of knowledge about the principles of LTA. Last but not least, the teacher also considered the possibility of facilitating 'pedagogical intelligence' (O'Donovan et al. 2008) in the GE class by inducting prospective teachers into the discourse of their future profession. Such early professionalization of learning had all the necessary prerequisites, since the students had already received training in pedagogy, pedagogical psychology, and the methods of teaching foreign languages.

4.2. Exploring LTA Course Delivery

The LTA course, under examination, addresses undergraduate students and has been conducted by the author of this article for four years. When the author first designed the LTA course, she ensured that the content is consistent with the students' future professional needs, learning motivation, and mental maturity. It is also consistent with the language assessment culture adopted in the community, especially in respect of the compatibility of assessment practices with institutional or individual ideologies, social expectations, and attitudes and values, as Inbar-Lourie (2008) states that such a course should be. Moreover, the course contributes to shaping critical awareness of the students' prospective assessment practices, following Scarino's views (2013).

The course content comprises the *stable knowledge base* (Brown and Bailey 2008), which includes: purposes of tests, types of tests, parameters of test usefulness, principles of test development, scoring, and feedback, as well as theory and practice of item/task writing for testing receptive skills. The course is implemented as a series of workshops that are built on the generic principles of workshop design, i.e., with the objectives for constructing new knowledge set at the beginning, technique of heuristic talk employed throughout the class, theoretical input combined with hands-on experience of doing a variety of tasks by learners in teams, and eventual statement of the outcomes of the workshop activities. It is worth mentioning that the workshop format is still considered innovative, as it is not practiced broadly in the local higher education.

The course is delivered in English, although it is substantially tailored to trainees' English proficiency by using simpler structures and more frequent lexis, resorting to explanation of terminology, making links to the LTA course delivered in L1, delivering at a slower pace, and translating new concepts into L1. Additionally, the course is supported by multiple and diverse handouts, short glossaries, and presentation slides, available before the class (on students' request). The formative assessment implemented includes teacher observation, questioning, quizzes, and items/tasks developed by teams. The summative assessment is comprised of the exit test and grades for the test tasks to test grammar, vocabulary, and reading, as developed by students individually. Profiling the content according to the tenets of task-based learning allows for covering the selected course content at a good pace within the allotted 14 contact hours. The course is completed with administering the exit test and Questionnaire 2.

Due to the pandemic of 2020–2021, the course was conducted remotely, which, to a certain extent, optimized its delivery. Sharing workshop slides enabled trainees to follow the oral input closely and in a focused, synchronous manner. It proved meaningful during

theoretical input, as well as when performing practical tasks. Additionally, interaction in breakout rooms allowed the trainees to trial items/tasks collaboratively, thus recreating the collegiality of real-life test developers' work.

Although some alterations did occur since the course was first described by Kvasova (2020), they pertained primarily to the teaching techniques (adaptations of the scenario), but not to the content of teaching. This fact allows for comparison of the learning outcomes achieved by the students of the current study (Set 1 of participants) with those achieved by the students in the previous research (Kvasova 2020) (Set 2 of participants).

4.3. Analysis

The descriptive data analysis method was employed to interpret the results of the questionnaires.

5. Results

5.1. Results of RQ1

Results for RQ1 are based on the responses to the Likert-type questions in Questionnaire 1.

Q.1–5 are focused on the conception of feedback (see Section 3.1). Q.1–5 aims to elicit what form(s) of feedback: (1) are most useful in promoting better learning; (2) are most useful for understanding positive and negative features of performance; (3) are most helpful in directing further work to particular aspects of performance; (4) match their self-assessment; (5) are most meaningful during Term 1, when GE course was taught. The students were to evaluate the forms of feedback (grades, oral/written comments, and suggestions) on the five typical activities (individual long turn, role play, revision translation, writing task, and test) as 'very helpful', 'quite helpful', 'not helpful', and 'not applicable' and assign from 3 to 0 points, respectively. The data entered in Table 1 are the mean values of the points assigned by the respondents to the five activities.

Table 1. Analysis of students' feedback received in the GE course.

No	Forms of Feedback That . . .	Numerical Grade	Comment	Suggestion	No Feedback
1	Promote progress in learning	11	17.8	10.3	0
2	Promote understanding of positive/negative aspects of performance	10.8	18	17	0
3	Help direct remedial work at particular aspects of performance	12.4	18	17.4	0
4	Coincide with self-assessment,	15	17	16.2	0
5	Were most meaningful in Term 1	13	17.8	17.2	0

The answers to Q.1–5 elicited the following. First, the "0" responses in the "No feedback" column suggest that the cohort of learners received feedback throughout the term and could distinguish receiving feedback from not receiving it. This suggests that they knew what feedback is. Second, the data show that the feedback was provided in at least three forms (numerical grade, comment, and suggestion), which suggests that the respondents were well-familiar with these forms of feedback and could make their judgements of them. Additionally, they indicated that the comment, closely followed by suggestion, appeared the most helpful and useful/clear form of feedback, which highly coincided with self-assessment and clearly promoted progress. The suggestion appeared more helpful, in terms of directing further work to particular aspects of the performance. The most valued forms of feedback, in terms of gauging improvement, were comment and suggestion.

Q.6–9 are centered around the conception of conscious knowledge of assessment construct and criteria (see Section 3.1). Q.6–9 aim to elicit to what extent/how frequently the students: (6) were made familiar with the focus of assessment in each of the activities; (7) were informed of the number and types of tasks included in the summative test and the scores they may get for them; (8) were informed of the criteria used to assess productive skills; (9) fully understood these criteria. The students were to assess the state of their familiarization with and/or the clarifications received about the written test construct and assessment criteria for productive skills (individual long turn, role plays, and writing tasks) using 'always, 'often', 'sometimes', or 'never' by scoring from 3–0 points, respectively. Table 2 presents the mean values of the points assigned per question by the respondents.

Table 2. State of students' knowledge about assessment criteria in the GE course.

No	State of Students'...	Always	Often	Sometimes	Never
6	Being familiarised with the focus of assessment in five activities	9	6.4	2.4	1
7	Being informed about tasks included in summative test and scores	9	9	0	0
8	Being informed of assessment criteria for productive skills	11.6	5.6	1	0
9	Being clear about assessment criteria for productive skills	12	5.6	0	0

The data show that the degree of respondents' familiarity with/knowledge about the construct and criteria was ensured in the majority of incidences. The total score of "sometimes" or "never" getting familiar with the assessment construct and criteria is not significant.

When asked Q.10 (to evaluate the effects of the knowledge of criteria on the learning produced), the respondents primarily mentioned an enhanced self-assessment and clarity about what needs to be improved, reflections about the quality of the assessed skills, a more conscious approach to doing tasks, and the ability to more consciously control their performance.

All in all, the data obtained via Questionnaire 1 suggest that, in the GE course, the students were exposed to varied forms of assessment. The most valued forms of feedback are associated with the ability to clearly realize what should be improved (and how), which is traditionally viewed as the most meaningful aspect of feedback in education. Similarly, the data demonstrate that the respondents are knowledgeable about assessment construct and criteria. Therefore, RQ1 "Did the trainees develop such CoA as (1) understanding of feedback and (2) knowledge of assessment construct and criteria during the 4-month GE course?" appears to be answered positively.

5.2. Results of RQ2

RQ2 was explored with the help of the exit test and Questionnaire 2.

5.2.1. Results of Exit Test

The exit test yielded an average of 90% correct answers, with the percentages ranging from 85 to 100, with a 90% median. Table 3 shows these results, compared with the results of the exit test taken by the comparison group (Kvasova 2020).

Table 3. Comparison of exit test scores in the current and prior studies.

Results of Study	No of Testees	Range of Scores	Mean	Median
Current study	19	85–100	90%	90%
Study of 2020	21	70–95	82%	85%

The scores are indicative of the effectiveness of the course, with those of the current study being considerably higher. This is attributed to better acquisition of the course content by the current cohort of students (Set 1). Thus, teaching and assessing by the book, in the previous term (Term 1), in the current study, had a positive effect on the learning outcomes in the LTA course.

5.2.2. Results of Questionnaire 2

RQ2 "Did the CoA developed within the GE course impact the students' LAL developed within the LTA course?" was also answered with data from the Likert-type and selected response questions. The descriptive analysis of these data showed the following.

Q1 consisted of 16 statements about the knowledge of particular concepts in the LTA course and development of practical skills formulated directly. These were viewed as self-evaluation of the LTA course learning outcomes by the respondents. These data, presented in Table 4, reflect the efficiency of learning the course content.

Table 4. Self-evaluation of LTA course content acquisition.

No	Statement	Strongly Agree	Agree	Unsure	Disagree	Strongly Disagree
1	I know the difference between formative and summative assessments	6	10	3	0	0
2	I know the difference between achievement and proficiency tests	7	6	6	0	0
3	I know what test validity is	3	7	9	0	9
4	I know what test reliability is	6	8	5	0	0
5	I know what test impact is	9	9	1	0	0
6	I know what test practicality is	9	10	0	0	0
7	I know what test transparency is	8	8	3	0	0
8	I know the components of a test task	11	6	2	0	0
9	I know what information a rubric to task should contain	11	6	2	0	0
10	I know the difference between selected and constructed responses	11	6	2	0	0
11	I know how a test is prepared, e.g., piloting	6	7	6	0	0
12	I know what test specification is and what it contains	5	7	7	0	0
13	I know various test formats (e.g., alternative choice, matching, gap-filling)	11	6	2	0	0
14	I can write tasks to test grammar	12	5	2	0	0
15	I can write tasks to test reading	12	5	2	0	0
16	I know/can use a rating scale to test speaking/writing	8	4	7	0	0
	Mean	8.4	6.8	3.6	0	0

The indexes vividly demonstrate the respondents' perceived strengths and weaknesses. They showed that the respondents were quite confident and "strongly agree" with the statements about their achievements, in an average of 8.4 of incidences, whereas there were no negative responses whatsoever. The highest score, 12, stands for item writing skills. The second highest one was 11 and shared by the responses related to test tasks and test

formats, while the indexes for some theory-related questions were moderate, ranging from 6 to 9. Unsurprisingly, the questions regarding knowledge about test specifications and test validity scored quite low, "5" and "3", respectively. It should be noted that a similar distribution of scores, across the theoretical content points, was received in the exit test, too.

In Q2, the statements invited students to evaluate the complexity of the LTA course and aspects that facilitated trainees' gains during the course. Amongst the 10 statements, those directly connected with the effects of pre-existing CoA are presented in Table 5. The answers showed that the majority of respondents expressed agreement with the respective statements, whereas no one expressed disagreement, which testifies to the essential contribution of studying the GE course to shaping CoA.

Table 5. Perceived impact of GE course experience on learning gains in LTA course.

No	Statement	Strongly Agree	Agree	Unsure	Disagree	Strongly Disagree
1	Acquisition of theoretical knowledge was made easier, thanks to taking the GE course in Term 1	8	8	3	0	0
2	Development of new practical skills was made easier, thanks to taking the GE course in Term 1	7	6	6	0	0

Q3, which is a selected response one, aims to specifically find out if the students had reflected on some major issues of LTA (such as feedback, validity, reliability, etc.) *during* or *prior* to taking the course in LTA.

The data entered in Table 6 suggest that the concepts of feedback and washback were primarily contemplated on by the respondents within the GE course. At that period of time, the respondents associated themselves with the learners, rather than the assessors. For the same reason, students were quite concerned about the clarity of rubrics and test transparency (Items 3 and 4). Assessment criteria together with the concepts of validity and reliability (Items 5–7) were contemplated on by the respondents, mostly while taking a profession-oriented LTA course. Obviously, this reflects the roles adopted by the respondents during the courses: the learner seeking feedback in the GE course, and the trainee mastering new concepts and skills in the LTA course.

The Likert-type Q4 focuses on the same concepts as in Q3, but the statements are formulated as beliefs/opinions of the major principles of LTA. The question intends to elicit the effects of learning experience within GE course on the acquisition of particular testing concepts within the LTA course. The most frequent highest degree of agreement was with the following four statements (see Table 7).

The statements formulated indirectly, i.e., in the form of beliefs, allowed for the detection of the respondents' internalization of the concepts of validity, transparency, and reliability prior to the LTA course in a tacit way. This is somewhat contradicting to the data obtained in Q3, in which the questions were formulated in a direct manner, i.e., using official language/terminology. This controversy may be attributed to different perspectives on the statements–beliefs and statements–directives, in which beliefs are naturally better understood by the respondents.

Table 6. Perceived points in time when initial reflections on LTA concepts occurred.

No	I Contemplated . . .	During GE Course	During LTA Course	Never
1	On meaningfulness of timely teacher feedback	12	9	0
2	How rubrics to test tasks are formulated	1	17	1
3	That unclear rubric to test task may impede performance	10	14	0
4	On meaningfulness of getting detailed information about forthcoming test paper	10	11	0
5	That breaches of test administration procedure may distort the results of testing	7	17	0
6	On criteria to assess my oral/written performance	2	15	0
7	How trustworthy the score awarded to my oral/written performance is	4	12	3
8	On the impact of fair/unfair grade on further studies	12	11	0

Table 7. Perceptions of GE course impact on acquisition of some testing concepts.

No	Statement	Strongly Agree	Agree	Unsure	Disagree	Strongly Disagree
1	Clear wording of the rubric to task promotes better performance	15	4	0	0	0
2	Students should know well how many and what task types are included in the test	14	3	2	0	0
3	Students should know the criteria of assessing speaking and writing	12	5	2	0	0
4	Feedback on performance should be timely	10	9	0	0	0

Q.5 aims to elicit respondents' perceptions about the effects of several academic courses and/or experiences on their learning gains in the LTA course.

As is seen in Table 8, the most essential impact was that of the GE course delivered in Term 1 (11 total positive responses), followed by prior reflections on assessment (10), additional reading (10), and TEFL course (7). The priority of the GE course testifies to the essential impact of the instructor's effort for repeated and systemic exposure of students to the tenets of language assessment in action. The credit paid to the TEFL course, on the one hand, testifies to the overall well-grounded approach that is adopted to teaching pedagogy in the university discussed. The high number of responses obtained by "prior reflections about assessment", testifying to students' critical awareness of assessment practices shaped before the formal training in LTA, is also noteworthy. Surprisingly, the respondents did not consider the part of the LTA course delivered in L1 or their prior experience of taking Independent School-Leaving Test as having positive effects on their learning.

The descriptive statistics suggests that RQ2 "Did the two CoA developed within the GE course impact the students' LAL developed within the LTA course?" was also resolved.

Table 8. Perceptions of impact of prior experience on the learning outcomes in the LTA course acquisition.

No	The Part of the LTA Course Delivered in English Is Internalized by Me Thanks to . . .	Strongly Agree	Agree	Unsure	Disagree	Strongly Disagree
1	The part of the LTA course delivered in L1	0	3	11	5	0
2	The TEFL course studied previously	0	7	9	3	0
3	The GE course studied in Term 1	2	9	6	2	0
4	Previous experience of taking standardized and other tests	1	4	3	11	0
5	My prior reflections about assessment	2	8	9	0	0
6	Additional readings	1	9	4	5	0
7	Help of other teachers	3	3	1	12	0

6. Discussion

The data of the study seem to vividly demonstrate that the cohort of students that had previously been taught the GE course by the assessment-minded instructor have coped quite well with the content of the LTA course. Comparing the results of their exit test with the results of Questionnaire 2, it seems that the students, through their perception of their mastery of the course content, confirmed the quantitative data obtained. When asked about the most positive impact of their prior knowledge and experiences, they largely mentioned the GE course. This fact confirms that the active, although tacit, engagement of students with assessment concepts and principles, through being taught and assessed during this course, appears effective. The assumption about the feasibility of conceptions of "understanding of feedback" and "knowledge of assessment construct and criteria" seems positive.

It is reasonable to now present the author's vision of the two differential types of LAL addressed in this study. The first is student initial AL, developed within the learners of the GE course. It may be defined as the understanding of assessment procedures, existing in a particular context, i.e., the ability to understand, interpret, and use feedback provided on one's performance, as well as the conscious knowledge of assessment criteria/standards enabling the control and evaluation of own performance. This LAL is developed in an active, though tacit, way via a systemic exposure to assessment procedures properly organised.

The emerging 'pedagogical intelligence', or rather interaction of the instructor with the future colleagues, suggests the idea of certain apprenticeship. Unlike 'folkways of learning' in 'apprenticeship of observation' by Lortie (1975), this relationship stems from the community of interests, mentor's willingness to share their teaching/assessing expertise, and mentee's aspiration to acquire the profession. The lessons and insights gleaned from these learning experiences may further affect the prospective teachers' identity and toolkit. However, special research is required to look into this issue if we want to discriminate between the mentor's influence and the influence of course of life, circumstances, and other factors that affect becoming a professional.

The other type of student LAL is the one developed via the course in LTA. This is the prospective teacher LAL, which, following Malone (2017), should not be confused with the LAL of a practicing teacher. In line with Scarino (2013) and Tsagari (2017), the LAL of prospective teachers is viewed as fully context-dependent. We may describe it in a

plethora of ways building on the well-known definitions of LAL, though with a significant amendment—the one regarding its relevance to, and consistence with, the curriculum of the course taught. The construct of this LAL will be collated with the curriculum and learner's background and maturity, as well as the purpose of their studying LTA.

For instance, Ukrayinska (2018), delineating the construct of classroom LAL of the English majors in their 5th year of study, provides a ramified system of skills and subskills, divided further into lower- and higher-level subskills. In the case of the LTA course aimed at second year Ukrainian majors, as was discussed throughout the article, the course content is concentrated and organized in a constructivist framework, with the language of the course delivery adapted to the learners' proficiency in English and techniques made as interactive as possible. For instance, when trained in a writing item to test language skills and reading, the students are engaged in peer-trialing and further modifications of tasks. Additionally, they are familiarized with assessing listening by doing test tasks of different formats themselves and further reporting on their experience and knowledge acquired. When it comes to the assessment of productive skills, the students are shown how to use rating scales and have an opportunity to assess original scripts/oral performances.

Both types of student LAL cannot be viewed outside the course content, institutional context, and, moreover, irrespective of individual perceptions and beliefs, as well as the predisposition towards learning and assessment, as parts of pedagogical work.

7. Limitations of the Study

The limitations of the current study are connected with the psychology-related issues of the attitudinal component of assessment competence, i.e., conceptions, beliefs, and attitudes. Certain steps have been made by scholars in assessing the attitudinal component of intercultural competence (e.g., Scarino 2009; Fantini 2010). Instruments of measuring CoA have been developed and efficiently used by Brown (2008), Brown et al. (2014), Opre (2015), and other scholars. However, given the significance of attitudes and beliefs for developing and examining LAL of various stakeholders, the tools of researching conceptions need further investigation, in terms of the creation of feasible methodology appropriate for small-scale classroom-based studies.

Another limitation refers to the research focus on a quite limited (4-month) timespan in each of the terms. As Deneen and Brown (2016) argue, CoA takes a longer time to shape/reshape than a typical academic course lasts. Therefore, a longitudinal study of conception-building processes may be a beneficial continuation of this research project in the future.

This small-scale study employed two questionnaires to elicit the students' perceptions, with questionnaires being marked for certain degree of subjectivity. Similarly, the classroom-based teacher action research was not devoid of a certain degree of subjectivity. The small sample size in the study did not provide the necessary scope of the data to use any methods to confirm its statistical significance, either.

8. Implications

There are certain implications of the current research. All assessment procedures within a GE course should be implemented, in compliance with the principles of good practice in LTA (ILTA 2020; EALTA 2006). Assessment principles should be strictly adhered to by the instructors, in order to enable students' repeated exposure to properly implemented assessment procedures, thus promoting building their tacit knowledge of assessment and shaping CoA.

Additionally, the integration of assessment in the course, with the regular engagement of students with assessment standards/criteria, should enable them to internalize the criteria and further monitor own learning. This should become a norm in teaching any language course by all instructors. Additionally, the creation of pedagogical intelligence (O'Donovan et al. 2008) in the teaching/learning environment should promote early specialization in

the prospective teacher education and, on the other hand, contribute to their induction to the profession.

As a result of the above-mentioned, the LAL of prospective teachers, as one of the objectives of language teacher education, may be efficiently developed.

9. Prospects for Further Research

The investigation of the impact of prior CoA on the efficiency of the LTA course implementation could be replicated in a larger-scale research, from a longitudinal perspective. In this case, learners' CoA could be considered in a more systematic way, including, in particular, the survey of the CoA developed during the pre-university period of learners' life. Additionally, the involvement of several teachers and researchers in the study could reduce the subjectivity of the study conducted by one person who has the dual role of teacher and researcher.

Finally, what is critical for future research is the development of reliable instruments to measure CoA and further interpret them properly.

10. Conclusions

The article presents the theoretical underpinnings and practical implementation of an action research into the impact of pre-existing CoA on the outcomes of learning in a stand-alone course in LTA. Based on the review of the relevant studies in the field and corroborated by the data obtained via Questionnaire 1, the two most significant CoA for this research context were found to be "understanding of feedback" and "conscious knowledge of assessment construct and criteria". This suggests the resolution of Research Question 1. The employment of Questionnaire 2 and results of the exit test, conducted at the end of the LTA course, enabled the detection of a positive impact of the CoA, developed within the GE course, in Term 1 on the acquisition of the fundamentals of LTA in Term 2. Thus, the RQ 2 is also viewed as answered.

To sum up, the findings suggest that the proper implementation of language assessment within a language course and creation of a professional, pedagogical discourse in the classroom may promote tangible positive learning outcomes in the LTA course that follows.

Appendix A
Questionnaire 1

Question	Numerical Grade	Comment	Suggestion	No Feedback
1 Which form of feedback promotes your better learning? *Put from 0 ("nothing") to 3 ("very much") as many times as necessary* Individual long turn Role play (paired) Written work (essay) Revision translation Mid-term test paper				
2 Which form of feedback is the most useful for you (helps to understand strengths and weaknesses of your performance)? *Put from 0 ("nothing") to 3 ("very useful")* Individual long turn Role play (paired)				

Question	Numerical Grade	Comment	Suggestion	No Feedback
Written work (essay) Revision translation Mid-term test paper				
3 Which feedback is more stimulating (helps to target further effort at particular aspects)? *Put from 0 ("nothing") to 3 ("very stimulating")* Individual long turn Role play (paired) Written work (essay) Revision translation Mid-term test paper				
4 Which feedback most often coincides with your expected score? *Put from 0 ("never") to 3 ("very often")* Individual long turn Role play (paired) Written work (essay) Revision translation Mid-term test paper				
5 Which feedback was the most important for you in GE classes in Term 1? *Put from 0 ("nothing") to 3 ("very important")* Individual long turn Role play (paired) Written work (essay) Revision translation Mid-term test paper				
6 Are you aware what is being assessed in each activity? *Put (+) as applies* Individual long turn Role play (paired) Written work (essay)	Fully aware	Quite aware	Not quite aware	Unaware
7 Do you receive information about number and types of tasks in the upcoming mid-term assessment? *Put (+) as applies* Individual long turn Role play (paired) Written work (essay)	Each time	Sometimes	Almost never	Never
8 Do you receive information about assessment criteria in the following tasks? *Put (+) as applies* Individual long turn Role play (paired) Written work (essay)	Each time	Sometimes	Almost never	Never
9 How well do you understand the assessment criteria? *Put (+) as applies* Individual long turn Role play (paired) Written work (essay)	Each time	Sometimes	Almost never	Never
10 What does understanding of assessment criteria help you to do? *Put (+) as applies* Consciously prepare for the performance Control your performance Assess yourself Set objectives for improvement Analyse your performance	Each time	Sometimes	Almost never	Never

Appendix B

Exit test

1 **Put the steps of developing a test (A–H) in the correct order (1–8). Write the number before the letter.**

_____ **A** Pre-testing/trialling
_____ **B** Pre-testing and revising the test
_____ **C** Familiarising with test specification
_____ **D** Revising the items
_____ **E** Writing items
_____ **F** Planning the test
_____ **G** Analysing the items
_____ **H** Receiving feedback from pre-testing

2 **Match the definitions (9–16) with the aspects of test usefulness (A–I). There is one aspect that you do not need. Write your answers next to numbers.**

A absence of bias   **D practicality**   **G transparency**
B authenticity   **E reliability**   **H validity**
C interactiveness   **F security**   **I washback/impact**

9 ____ A good classroom test should be teacher-friendly. A teacher should be able to develop it within the available time and with available resources.

10 ____ Good testing or assessment strives to use formats and tasks that mirror the real-world situations and contexts in which students would use the target language.

11 ____ If we want to use the results of a test to take a decision, we must be confident that the results give us information that is relevant to our decision. The test must measure what we think it measures – and that measurement must be a good basis for the decision we make.

12 ____ If students perceive that the tests are markers of their progress toward achieving clear course outcomes, they have a sense of accomplishment. "Test-driven" curricula and only learning "what they need to know for the test" are said to make learning less efficient.

13 ____ Clear and accurate information about testing given to students should include outcomes to be evaluated, formats used, weighting of items and sections, time allowed to complete the test, and grading criteria.

14 ____ It should not matter when we give a test. If we give it on a Monday, it should give the same results as if we gave the test on a Saturday. It should not matter who scores the test. If one teacher scores a test as 7 correct out of 10, anybody else scoring the test should arrive at the same score.

15 ____ The test clearly corresponds to test takers' age and actual interests, and the language used in the questions and instructions is appropriate for their level.

16 ____ If we want to use a test not once, we should make sure that the students do not share the answers that they think are correct with those who will write the test after them. Nor should we overlook cheating during the test administration.

3 Match test formats (17–25) with rubrics (A–J). There is one rubric that you do not need. Write your answers next to numbers.

17 ____ Matching jumbled questions to parts of an interview.
18 ____ Filling gaps in text with clauses, sentences, parts of text
19 ____ Banked gap-filling
20 ____ Short-answer questions
21 ____ Matching short texts/parts of text and headings
22 ____ Sentence transformation/paraphrase
23 ____ Alternative choice
24 ____ Multiple choice
25 ____ Matching texts and questions

A Read the text. Choose the most suitable heading from the list (A- . . .) for each part (1- . . .) of the text. There is one extra heading, which doesn't match any part.
B Fill in the gaps in the text with the correct words/phrases (A, B, C, or D).
C Read the text. Answer questions (1- . . .) below using a maximum of THREE words.
D Read the text below. Match parts of the text (A- . . .) to questions (1- . . .). There is/are . . . extra question(s), which does not/do not match any part.
E Read the text below. Some sentences have been removed from the text. Choose from sentences (A- . . .) the one which best fits each of (1- . . .). There is one extra sentence, which you do not need to use
F Fill in the gaps (1- . . .) in the text by correct verb forms. Write the words in the answer grid.
G Readan interview with . . . . The questions have been mixed up. Match the interviewer's questions (A- . . .) below to answers (1- . . .). There is one extra question, which doesn't match any answer.
H Fill in the gaps in the text below. Choose the most appropriate word from the list (A- . . .) for each gap (1- . . .) in the text.
I Fill in gaps in sentences (1- . . .) so that they keep the same meaning.Use the words in brackets.
J Choose a correct alternative *a* or *b* to complete sentences (1- . . .).

References

Anam, Syafi'ul, and Nanin V. W. Putri. 2021. How Literate am I about Assessment: Evidence from Indonesian EFL Pre-service and In-service Teachers. *English Review: Journal of English Education* 9: 151–62. [CrossRef]

Barnes, Nicole, Helenrose Fives, and Charity M. Dacey. 2014. Teachers' Beliefs about Assessment. In *International Handbook of Research on Teachers' Beliefs*. Edited by Helerose Fives and Michele G. Gill. London: Routledge, pp. 284–300. ISBN 9780415539258.

Beziat, Tara L. R., and Bridget K. Coleman. 2015. Classroom Assessment Literacy: Evaluating Pre-service Teachers. *The Researcher* 27: 25–30.

Bøhn, Henrik, and Dina Tsagari. 2021. Teacher Educators' Conceptions of Language Assessment Literacy in Norway. *Journal of Language Teaching and Research* 12: 222–33. [CrossRef]

Boyd, Ashley, Jennifer Jones Gorham, Julie Ellison Justice, and Janice L. Anderson. 2013. Examining the Apprenticeship of Observation with Preservice Teachers: The Practice of Blogging to Facilitate Autobiographical Reflection and Critique. *Teacher Education* 40: 27–49.

Brookhart, Susan M. 2011. Educational Assessment Knowledge and Skills for Teachers. *Educational Measurement: Issues and Practice* 30: 3–12. [CrossRef]

Brown, Gavin T. L. 2008. *Conceptions of Assessment: Understanding What Assessment Means to Teachers and Students*. New York: Nova Science Publishers.

Brown, James Dean, and Kathleen M. Bailey. 2008. Language Testing Courses: What are They in 2007? *Language Testing* 25: 349–83. [CrossRef]

Brown, Gavin T. L., and Ana Remesal. 2012. Prospective Teachers' Conceptions of Assessment: A Cross-Cultural Comparison. *The Spanish Journal of Psychology* 15: 75–89. [CrossRef]

Brown, Gavin T. L., Reza Pishghadam, and Shaghayegh S. Sadafian. 2014. Iranian University Students' Conceptions of Assessment: Using Assessment to Self-improve. *Assessment Matters* 6: 5–33. [CrossRef]

Coombe, Christine, Keith Folse, and Nancy Hubley. 2007. *A Practical Guide to Assessing English Language Learners*. Ann Arbor: University of Michigan Press.

Coombe, Christine, Hossein Vafadar, and Hassan Mohebbi. 2020. Language Assessment Literacy: What Do We Need to Learn, Unlearn, and Relearn? *Language Testing in Asia* 10: 1–16. [CrossRef]

Cox, Stephanie E. 2014. Perceptions and Influences Behind Teaching Practices: Do Teachers Teach as They Were Taught? Master's Science Thesis, BrighamYoung University, Provo, UT, USA.

Crossman, Joanna. 2007. The Role of Relationships and Emotions in Student Perceptions of Learning and Assessment. *Higher Education Research and Development* 26: 313–27. [CrossRef]

Davies, Alan. 2008. Textbook Trends in Teaching Language Testing. *Language Testing* 25: 327–48. [CrossRef]

DeLuca, Christopher, Daniele Lapointe-McEwan, and Ulemu Luhanga. 2016. Teacher Assessment Literacy: A Review of International Standards and Measures. *Educational Assessment, Evaluation and Accountability* 28: 251–72. [CrossRef]

Deneen, Christopher C., and Gavin T. L. Brown. 2011. The Persistence of Vision: An Analysis of Continuity and Change in Conceptions of Assessment within a Teacher Education Program. Paper presented at the 37th Annual Conference of the International Association of Educational Assessment (IAEA 2011), Mandaluyong City, Philippines, October 23–28.

Deneen, Christopher C., and Gavin T. L. Brown. 2016. The Impact of Conceptions of Assessment on Assessment Literacy in a Teacher Education Program. *Cogent Education* 3: 1225380. [CrossRef]

European Association for Language Testing and Assessment (EALTA). 2006. EALTA Guidelines for Good Practice in Language Testing and Assessment. Available online: https://www.ealta.eu.org/guidelines.htm (accessed on 10 June 2010).

Evans, Carol. 2013. Making Sense of Assessment Feedback in Higher Education. *Review of Educational Research* 83: 70–120. [CrossRef]

Fantini, Alvino E. 2010. Assessing Intercultural Competence: Issue and Tools. In *The SAGE Handbook of Intercultural Competence*. Edited by Darla K. Deardorff. Thousand Oaks: SAGE Publications, Inc., pp. 456–67. ISBN 978-1-4129-6045-8.

Fröjdendahl, Birgitta. 2018. Pre- and In-service Teachers' Assessment Literacy—A Qualitative Approach. *Literacy Information and Computer Education Journal (LICEJ)* 9: 2886–94. [CrossRef]

Fulcher, Glen. 2012. Assessment Literacy for the Language Classroom. *Language Assessment Quarterly* 9: 113–32. [CrossRef]

Giraldo, Frank. 2018. Language Assessment Literacy: Implications for Language Teachers. *Profile: Issues in Teachers' Professional Development* 20: 179–95. [CrossRef]

Giraldo, Frank. 2019. Language Assessment Practices and Beliefs: Implications forLanguage Assessment Literacy. *HOW* 26: 35–61. [CrossRef]

Giraldo, Frank, and Daniel Murcia. 2019. Language Assessment Literacy and the Professional Development of Pre-Service Language Teachers. *Colombian Applied Linguistics Journal* 21: 243–59. [CrossRef]

Grainger, Peter R., and Lenore Adie. 2014. How do Preservice Teacher Education Students Move from Novice to Expert Assessors? *Australian Journal of TeacherEducation* 39: 89–105. [CrossRef]

Green, Anthony. 2014. *Exploring Language Assessment and Testing. Language in Action*. New York: Routledge.

Hasselgreen, Angela, Cecile Carlsen, and Hildegunn Helness. 2004. European Survey of Language and Assessment Needs. Part One: General Findings. Available online: http://www.ealta.eu.org/documents/resources/survey-report-pt1.pdf (accessed on 23 April 2011).

Herrera Mosquera, Leonardo, and Osmar David González Meléndez. 2021. Perceptions and Beliefs of EFL Students Regarding Classroom Assessment at a Colombian University. *Folios* 54: 169–86. [CrossRef]

Higher Education Academy Project Report. 2012. A Marked Improvement: Transforming Assessment in Higher Education. Available online: https://www.heacademy.ac.uk/system/files/A_Marked_Improvement.pdf (accessed on 14 June 2021).

Hildén, Raili, and Birgitta Fröjdendahl. 2018. The Dawn of Assessment Literacy—Exploring the Conceptions of Finnish Student Teachers in Foreign Languages. *Apples: Journal of Applied Language Studies* 12: 1–24. [CrossRef]

Hill, Mary F., and Gayle Eyers. 2016. Moving from Student to Teacher: Changing Perspectives about Assessment through Teacher Education. In *Handbook of Human and Social Conditions in Assessment*. Edited by Gavin T. L. Brown and Lois R. Harris. New York: Routledge, pp. 103–28. ISBN 9781138811553.

Hutchings, Pat. 2005. Building Pedagogical Intelligence. Carnegie Perspectives. Available online: http://www.carnegiefoundation.org/perspectives/ (accessed on 28 July 2005).

Inbar-Lourie, Ofra. 2008. Constructing a Language Assessment Knowledge Base: A Focus on Language Assessment Courses. *Language Testing* 25: 385–402. [CrossRef]

Inbar-Lourie, Ofra. 2013. Guest Editorial to the Special Issue on Language Assessment Literacy. *Language Testing* 30: 301–7. [CrossRef]

International Language Testing Association (ILTA). 2020. Guidelines for Practice. Available online: https://www.iltaonline.com/page/ILTAGuidelinesforPractice (accessed on 3 May 2015).

Knapp, Nancy Flanagan. 2012. Reflective journals: Making constructive use of the "apprenticeship of observation" in preservice teacher education. *Teaching Education* 23: 323–40. [CrossRef]

Kremmel, Benjamin, and Luke Harding. 2019. Towards a Comprehensive, Empirical Model of Language Assessment Literacy across Stakeholder Groups: Developing the Language Assessment Literacy Survey. *Language Assessment Quarterly* 1: 100–20. [CrossRef]

Kvasova, Olga. 2009. *Fundamentals of Language Testing*. Kyiv: Lenvit. (In Ukrainian)

Kvasova, Olga. 2020. Will a Boom Lead to a Bloom? Or How to Secure a Launch of Language Assessment Literacy in Ukraine. In *Language Assessment Literacy: From Theory to Practice*. Edited by D. Tsagari. Newcastle: Cambridge Scholars Publishing, pp. 116–36.

Kvasova, Olga, and Tamara Kavytska. 2014. The Asessment Competence of University Foreign Language Teachers: A Ukrainian Perspective. *Language Learning in Higher Education* 4: 159–77. [CrossRef]

Lam, Ricky. 2015. Language Assessment Training in Hong Kong: Implications for Language Assessment Literacy. *Language Testing* 32: 169–97. [CrossRef]

Looney, Anne, Joy Cumming, Fabienne van Der Kleij, and Karen Harris. 2018. Reconceptualising the Role of Teachers as Assessors: Teacher Assessment Identity. *Assessment in Education: Principles, Policy & Practice* 25: 442–67. [CrossRef]

Lortie, Dan C. 1975. *Schoolteacher: A Sociological Study*. London: University of Chicago Press.

Lutovac, Sonja, and Maria A. Flores. 2021. Conceptions of Assessment in Pre-service Teachers' Narratives of Students' Failure. *Cambridge Journal of Education* 52: 55–71. [CrossRef]

Malone, Margaret E. 2017. Including Student Perspectives in Language Assessment Literacy. Paper presented at 39th Language Testing Research Colloquium, Bogotá, Colombia, July 17–21; p. 83.

Martínez Marín, Juan Diego, and Maria Camila Mejía Vélez. 2021. Master of TESOL Students' Conceptions of Assessment: Questioning Beliefs. *Revista Actualidades Investigativas en Educación* 21: 1–29. [CrossRef]

Massey, Kyle D., Christopher DeLuca, and Danielle LaPointe-McEwan. 2020. Assessment Literacy in College Teaching: Empirical Evidence on the Role and Effectiveness of a Faculty Training Course. *To Improve the Academy* 39: 209–38. [CrossRef]

Noble, Christy. 2021. Assessment Literacy for Students and Staff. Available online: https://medicine.uq.edu.au/faculty-medicine-intranet/md-program-staff-development/supporting-student-learning/assessment-literacy-students-and-staff (accessed on 6 August 2021).

O'Donovan, Berry, Margaret E. Price, and Chris Rust. 2004. Know What I Mean? Enhancing Student Understanding of Assessment Standards and Criteria. *Teaching in Higher Education* 9: 325–35. [CrossRef]

O'Donovan, Berry, Margaret E. Price, and Chris Rust. 2008. Developing Student Understanding of Assessment Standards: A Nested Hierarchy of Approaches. *Teaching in Higher Education* 13: 205–17. [CrossRef]

Odo, Dennis Murphy. 2016. An Investigation of the Development of Pre-service Teacher Assessment Literacy through Individualized Tutoring and Peer Debriefing. *Journal of Inquiry & Action in Education* 7: 31–61.

Oleson, Amanda, and Matthew T. Hora. 2013. Teaching the Way They Were Taught? Revisiting the Sources of Teaching Knowledge and the Role of Prior Experience in Shaping Faculty Teaching Practices. *Higher Education* 68: 29–45. [CrossRef]

Opre, Dana. 2015. Teachers' Conceptions of Assessment. *Procedia-Social and Behavioral Sciences* 209: 229–33. [CrossRef]

Pill, John, and Luke Harding. 2013. Defining the Language Assessment Literacy Gap: Evidence from a Parliamentary Inquiry. *Language Testing* 30: 381–402. [CrossRef]

Scarino, Angela. 2009. Assessing Intercultural Capability in Learning Languages: Some Issues and Considerations. *Language Teaching* 42: 67–80. [CrossRef]

Scarino, Angela. 2013. Language Assessment Literacy as Self-awareness: Understanding the Role of Interpretation in Assessment and in Teacher Learning. *Language Testing* 30: 309–27. [CrossRef]

Smith, Calvin D., Kate Worsfold, Lynda Davies, Ron Fisher, and Ruth McPhail. 2013. Assessment Literacy and Student Learning: The Case for Explicitly Developing Students 'Assessment Literacy. *Assessment & Evaluation in Higher Education* 38: 44–60. [CrossRef]

Smith, Lisa F., Mary F. Hill, Bronwen Cowie, and Alison Gilmore. 2014. Preparing Teachers to Use the Enabling Power of Assessment. In *Designing Assessment for Quality Learning. The Enabling Power of Assessment*. Edited by Claire Wyatt-Smith, Klenowski Valentina and Peta Colbert. Dordrecht: Springer Science + Business Media, pp. 303–23. ISBN 978-94-007-5902-2.

Solnyshkina, Marina I., Elena N. Solovova, Elena V. Harkova, and Aleksander S. Kiselnikov. 2016. Language Assessment Course: Structure, Delivery and Learning Outcomes. *International Journal of Environmental & Science Education* 11: 1223–29. [CrossRef]

Stiggins, Richard J. 1999. Evaluating Classroom Assessment Training in Teacher Education Programs. *Educational Measurement: Issues and Practice* 18: 23–7. [CrossRef]

Taylor, Lynda. 2013. Communicating the Theory, Practice and Principles of Language Testing to Test Stakeholders: Some Reflections. *Language Testing* 30: 403–12. [CrossRef]

Tsagari, Dina, and Karin Vogt. 2017. Assessment Literacy of Foreign Language Teachers around Europe: Research, Challenges and Future Prospects. *Papers in Language Testing and Assessment* 6: 41–63.

Tsagari, Dina. 2017. The Importance of Contextualizing Language Assessment Literacy. Paper presented at 39th Language Testing Research Colloquium, Bogotá, Colombia, July 17–21; pp. 82–3.

Tsagari, Dina, Karin Vogt, Veronica Froelich, Ildiko Csépes, Adrienn Fekete, Anthony Green, Liz Hamp-Lyons, Nicos Sifakis, and Stefania Kordia. 2018. Handbook of Assessment for Language Teachers. Available online: http://taleproject.eu/ (accessed on 18 May 2018).

Tuchyna, Natalia, Inna Zharkovska, Nina Zaitseva, Ihor Kamynin, Maryna Kolesnik, Lada Krasovytska, Halyna Krivchykova, and Tetiana Merkulova. 2019. *A Way to Success: English for University Students. Year 1. Student's Book*, 2nd ed. Kharkiv: Folio, ISBN 978-966-03-7136-1.

Ukrayinska, Olga. 2018. Developing Student Teachers' Classroom Assessment Literacy: The Ukrainian Context. In *Revisiting the Assessment of Second Language Abilities: From Theory to Practice. Second Language Learning and Teaching*. Edited by Sahbi Hidri. Cham: Springer International Publishing, pp. 351–71. ISBN 978-3-319-62884-4. [CrossRef]

Vogt, Karin, Dina Tsagari, and Georgios Spanoudis. 2020. What do Teachers Think They Want? A Comparative Study of In-Service Language. Teachers' Beliefs on LAL TrainingNeeds. *Language Assessment Quarterly* 17: 386–409. [CrossRef]

Vogt, Karin, and Dina Tsagari. 2014. Assessment Literacy of Foreign Language Teachers: Findings of a European Study. *Language Assessment Quarterly* 11: 374–402. [CrossRef]

Volante, Louis, and Xavier Fazio. 2007. Exploring Teacher Candidates' Assessment Literacy: Implications for Teacher Education Reform and Professional Development. *Canadian Journal of Education* 30: 749–70. [CrossRef]

Wach, Aleksandra. 2012. Classroom-based Language Efficiency Assessment: A Challenge for EFL Teachers. In *Glottodidactica XXXIX*. Poznań: Adam Mickiewicz University Press.

Xu, Yueting, and Gavin T. L. Brown. 2016. Teacher Assessment Literacy in Practice: A Reconceptualization. *Teaching and Teacher Education* 58: 149–62. [CrossRef]

Xu, Yueting, and Liyi He. 2019. How Pre-service Teachers' Conceptions of Assessment Change Over Practicum: Implications for Teacher Assessment Literacy. *Frontiers in Education* 4: 145. [CrossRef]

Yetkin, Ramazan. 2018. Exploring Prospective Teachers' Conceptions of Assessment in Turkish Context. *European Journal of Education Studies* 4: 133–47. [CrossRef]