

Language assessment literacy of language teachers in the context of adult education in Spain

Julia Zabala Delgado, Universitat Politècnica de València

Cristina Rodriguez, EOI de Santiago de Compostela

Recent years have seen a growth in the use of language tests as measures of accountability within the education sector. In many countries, governmental institutions have promoted the involvement of teachers in language testing, providing training to boost the language assessment literacy (LAL) of teachers. This study aims to analyse the results of a large-scale effort to increase teachers' LAL within the context of public language education for adults in Spain by shedding light on the scale and nature of teacher LAL, the impact of training as perceived by teachers, and their self-perceived further needs. Results show that, similarly to other countries in which teacher LAL has been studied, training in assessment is strongly influenced by contextual factors. Moreover, teachers perceive that this training has an impact not only on assessment-related tasks, but on their general teaching practice. Lastly, the findings reveal a significant correlation between the contents of assessment training courses and the teachers' perception of further training needs. This could indicate that the more teachers learn about specific areas of language assessment, the more training in assessment they feel is needed, suggesting a gap in teachers' awareness of their own LAL that materialises once training is provided.

Key words: language assessment literacy, language testing, teacher training

Introduction

Teachers can play an important role in the formative uses of classroom-based assessment to improve teaching and learning. They can also offer a critical analysis of the adequacy of standardised tests for different purposes. As key stakeholders, teachers have a professional responsibility to use assessment to improve learning. However, as in other European countries (Hasselgreen et al.,

Email address for correspondence: crodriguez@eoisantiago.org

© The Author(s) 2022. This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International License, which permits the user to copy, distribute, and transmit the work provided that the original authors and source are credited. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

2004), training in language assessment has not been extensive in teacher education in Spain and until recently was limited to professional development courses that focused mostly on practical implementation of classroom-based testing. In the context of adult language education in Spain, teachers' involvement in language assessment has been centred on classroom-based assessment and assessment to support learning. Proficiency assessment for public accountability has traditionally been centralised and designed by external experts appointed by the educational authorities.

The publication of the Common European Framework of Reference for Languages and its Companion Volume (Council of Europe, 2001, 2020) as a tool for harmonisation and the implementation of language policies promoted by the Council of Europe encouraged Spanish educational authorities, as well as other European countries, to introduce formal examinations as measures of accountability for language programmes. The use of high-stakes examinations in educational contexts implies the involvement of a wider range of stakeholders in the interpretation and use of tests to promote greater fairness (Deygers & Malone, 2019; Pill & Harding, 2013; Rea-Dickins, 1997; Taylor, 2013) and the need to take local contexts into account in the organisation of LAL training (Coombe et al., 2020; Giraldo, 2020; Yan & Fan, 2020). Therefore, the need has grown for teachers involved in language programmes to increase their language assessment literacy (LAL), which we define as their understanding of the principles and practice required to design and use assessment tasks and the impact of these practices on the wider community. However, because teachers are amongst the stakeholders with a more complex role in the testing process (Alderson et al., 1995), administrations and higher education institutions cannot successfully implement policy changes without previously training their teachers in the specific characteristics of standardised high-stakes language examinations. Consequently, an important effort was made on the part of Spanish educational institutions to assign resources to organise training programmes in standardised assessment for teachers. However, organising such LAL training needs to include an understanding of the contrasting paradigms of psychometric testing and the sociocultural views of learning of classroom-based formative assessment (Scarino, 2013). Institutions have to acknowledge that their teachers find themselves caught between the compulsory institutional requirements of a psychometric approach and their general teaching practice, in which classroom-based assessment played a prominent role.

For LAL to have an impact on teacher practice, these training programmes should be constructed in a way that allows teachers to relate to their contents, to regard them as positive and practical, and accept them as important for their teaching profession (Vogt et al., 2020). LAL should be in this sense expanded to include teachers' conceptions of their role as assessors, as defined by Brown

(2011), which will ultimately influence the way in which teachers understand and use language testing. The need to broaden this kind of professional education and adapt it to comply with the needs of language teachers calls for an understanding of the breadth and content of LAL training and of teachers' perceptions of its impact and usefulness. This study aims to explore these themes within the Spanish context, in which little is known of the extent and qualities of training in assessment for teachers of adult education, by use of a questionnaire in which the characteristics and views of teachers were collected and analysed.

Background to the study

The accreditation of language competence by means of formal examinations for adults studying foreign languages has in recent years become a common procedure in Europe as a result of changes in the labour market and legislative changes in education. Language competence has become a requisite for university graduation in many countries, following the guidelines set by the Bologna Declaration on encouraging plurilingualism in Higher Education (Bologna Declaration, 1999). In the case of Spain, government-regulated language education outside the school system is organised by two large organisations: (1) Escuelas Oficiales de Idiomas (EOIs), a network of public language schools for adults, and (2) University Language Centres within the Spanish Association of Language Centres in Higher Education (ACLES). Both organisations are regulated by the national and regional ministries of education and their institutions provide foreign and second language classes for the general public in the first case and the university sector in the second, following a government-regulated curriculum. These institutions have a widespread presence throughout Spain, with 321 EOI schools and 65 language centres in higher education institutions, according to the latest government figures (<https://www.educacionyfp.gob.es/servicios-al-ciudadano/estadisticas.html>).

Additionally, both institutions are involved in standardised language assessment as part of accountability systems that were created to ensure the quality of educational standards. As government-regulated institutions, they have an added social responsibility when acting as examination boards and are perceived by the public as granters of governmental approval.

These institutions administer standardised, high-stakes proficiency tests for adults in different foreign languages, with a target test taker population over 16 years of age that includes both EOI students and external candidates. These general proficiency examinations are not compulsory for the courses, can be taken independently from them, and lead to the award of certificates that can be used for graduation requirements, academic and professional mobility, civil servant merits, and access to the labour market. As proficiency examinations,

they are linked to the CEFR levels and consist of a minimum of four parts (reading, writing, listening and speaking). Originally, teachers in both institutions were responsible for teaching and classroom assessment, and standardised proficiency tests were developed by groups of experts. However, changes in governmental policies made teachers responsible for the development of proficiency tests including test specifications, task design, standardisation, and linking to the CEFR, as well as test administration, rating and communication of results. Consequently, an important institutional effort has been made to offer adequate training for teachers in their newfound responsibilities, which were previously seen as separate from their general teaching practice, that is, from language teaching and classroom-based formative assessment.

Accreditation of language competence in EOI and ACLES centres in Spain underwent a change between 2007 and 2008, with the implementation of new requirements that demanded changes in exam formats and standardisation efforts to comply with a common framework for exam development. Language teachers were to be involved in the development of standardised exams that were separate and had different specifications to the assessment practices carried out in the courses. Furthermore, exam development and specifications for these proficiency exams were modified to comply with international quality standards and make results transferable to other contexts, national or international. Additionally, quality control procedures were introduced pre- and post-administration, requiring teachers to undergo further training and take on additional exam development responsibilities within their institutions. Since classroom-based assessment was at the core of their training as language teachers and identified as part of their teaching practice, the language assessment training that was provided focused on the skills, knowledge, and understanding of assessment within the context of standardised high-stakes language tests. This included formal courses, participation in rating standardisation events, and other types of training events such as seminars and conferences. These types of formal assessment training events for teachers have increased in number in recent years in the context of adult language education in Spain (Rodríguez, 2015), and the language assessment literacy of the teachers involved in these processes has come into focus as a crucial element for the validity of the assessments carried out by these institutions when acting as examination boards.

Literature Review

Language assessment literacy for teachers

The role of language teachers becomes clear in the context of language assessment as a broad multifaceted term. They can be seen as the agents that

bring assessment closer to other stakeholders and also the facilitators for assessment to be used appropriately in different situational contexts, not only in terms of assessment for learning, but also in terms of critical understanding. However, encouraging LAL amongst language teachers involves bringing scholarship closer to their everyday practice. In fact, one of the main problems to overcome when encouraging LAL is the view that it is a field that is foreign to language teaching (Alderson, 2001; Davies, 2008; Fulcher, 2012; Pill & Harding, 2013; Taylor, 2009) and as noted by Stabler-Havener (2018), some of the most influential definitions of LAL in the literature (Fulcher, 2012; O'Loughlin, 2013) fail to include the word "language" and refer to "skills" or "abilities". Teachers have been given the responsibility of implementing educational policies and harmonising language teaching and assessment in the classroom and are presented with a conflicting situation in which their teacher role clashes with their assessor role. As a consequence, teachers may come to see assessment as irrelevant to their teaching and the teaching and learning process (Brown, 2011).

Many studies (cited below) have focused on exploring teachers' views on assessment, often with the intention of determining how best to overcome this perceived potential clash.

A preferred approach has been to study teachers themselves, making them a part of the process by analysing their self-reported level of assessment literacy, their needs and their perceptions of tests and testing. In the European context, Hasselgreen et al. (2004) conducted a European survey under the auspices of ENLTA, the European Network for Language Testing and Assessment, to find out about the needs of three groups: language teachers, language teacher trainers and what they referred to as 'experts', that is, professionals employed by organisations designing examinations. In their findings, they describe the needs of teachers as focussed on alternative forms of assessment (portfolios, assessing intercultural skills, informal assessment and self-assessment based on feedback as well as on the use of statistics to ensure the validity and reliability of the results. Huhta et al. (2005) carried out part two of the ENLTA survey focusing on language teachers and presenting results on European regions and countries, with results following a similar direction. The survey was again used by Vogt and Tsagari (2014), who adapted it to define the target group and focused only on language teachers in seven European countries with no assessment roles. Their findings corroborated Hasselgreen et al.'s to an extent, but they also identified a need for training in the assessment of the "four skills" and grammar and vocabulary, which had seemed less relevant in the original study. Fulcher (2012) also developed a questionnaire that focused on language teachers, recommending a balance between classroom and normative assessment when designing LAL courses. Similar attempts were made in other contexts using the Classroom Assessment Literacy Inventory (CALI). Mertler (2009) carried out a

study in the United States and reported the benefits of LAL training permeating into their work in the classroom. Xu and Brown (2017) used the inventory in China and corroborated teachers' need for LAL as well as the importance of the training to be contextualised.

Research into LAL needs has not only been attempted bottom-up by asking stakeholders directly, but also top-down by analysing materials used for training (Bailey & Brown, 1996; Brown & Bailey, 2008; Davies, 2008). However, none of these studies have investigated the needs of language teachers with responsibilities in high-stakes language testing. In this sense, the present study differs from the study conducted by Hasselgreen et al. (2004) in that it focuses specifically on language teachers and, from that carried out by Huhta et al. (2005), Fulcher (2012) and Vogt and Tsagari (2014), in that it focuses on language teachers with responsibilities in high-stakes language testing. What remains clear is that, although teachers have been prioritised as a key element, this process has not been without obstacles, as their twofold role often requires them to deal with diverging priorities.

Challenges in promoting assessment literacy for teachers

There is general agreement that a strong level of assessment literacy is a vital part of teacher professionalism (Xu & Brown, 2016), but much of the research focusing specifically on the assessment literacy of language teachers (Hasselgreen et al., 2004; Jin, 2010; Lam 2015; Malone, 2008; Vogt & Tsagari, 2014) has shown that it remains marginalized in teacher education programmes (Lam, 2015). The perception of the field as separate from second language acquisition and the psychometric component inherent in measurement activities have alienated many teachers. In fact, as stated by Popham (2006), teachers and administrators perceive that the core construct of training modules in education measurement has "little to do with the realities of the classroom" (p. 84) and requires further professional development to supply teachers with the assessment knowledge they need. Ultimately, if tests are to be more than policy instruments, teachers need to become involved to help create tests that are closer not only to the curriculum but also to the classroom where this curriculum is put into practice.

Yet making the field of assessment more accessible to teachers has proven difficult, in terms of not only ending misconceptions about the field, but also overcoming practical issues such as time constraints or availability of resources. After years of understanding assessment as "testing" or as standardised testing, teachers and educational institutions need to consider assessment as a key component of the process that feeds from and into the classroom. As previously mentioned, assessment needs to be understood as a tool to provide the teaching profession with a role in the shaping of educational policies. Furthermore, this

process is not one-sided, and the field of language testing can open its doors to new perspectives and different agents, both internal and external to the process of teaching and learning. Despite this, and as stated by Scarino (2013) "assessment remains the aspect of the curriculum and teaching and learning practices that is least amenable to change" (p. 310), as a result of deep-rooted traditions and the institutional nature of assessment acting as impediments to reflection in many contexts. To change assessment through its agents, research needs to be focused on their conceptions as defined by Brown (2011), which are rooted in their previous experiences and beliefs.

Research Focus

The aim of this study was to examine what kind of LAL training Spanish language teachers have received, particularly regarding standardised testing, the perceived impact this training has had on their assessment practices, and their perceived future needs. The focus was directed towards language teachers in adult education. Although efforts have been made to train teachers in the development and correct use of tests by using the findings of the studies in the literature, little is known about the perceived impact of this training on teachers. Consequently, our study aimed to answer the following research questions:

1. Have the teachers working in these institutions had specific training in standardised language assessment and standardisation of criteria, and if so, what was the length, nature and contents of such training?
2. What impact - understood as positive impact - do the teachers perceive the contents of this training to have had on their work?
3. What are their self-perceived future needs in language assessment?

Methodology

In order to address these questions a cross-sectional quantitative exploratory study was designed. A questionnaire was developed to ask teachers about their training in language assessment and their participation in standardisation sessions (RQ1), the content of the training and their beliefs about the impact it had had on different aspects of their work (RQ2) and, their self-perceived needs as regards further training (RQ3). The questionnaire was trialled with four EOI teachers and two University Language Centre teachers, representative of the two target populations, to determine that the instructions and questions were clear and understandable, and that suitable information prompts had been included. The questionnaire was created using Google Forms to allow for wide

dissemination and was designed to allow for anonymity, as many of the questions probed the respondents' own professional training and background.

The participants in the study were language teachers working for the two organisations and involved in standardised assessment to various degrees. The respondents comprised teachers of over seven different languages, although over 70% were English teachers, the majority of whom had more than five years' experience in teaching.

The online questionnaire (see Appendix) was sent to the 11 EOI schools in the autonomous community of Galicia and 61 ACLES university language centres across Spain. A total of 307 responses were collected from teachers working in these two institutions (EOI, n=114; ACLES, n=193). The resulting data was analysed using SPSS for descriptive statistics and the results were compared across the two contexts. The Cronbach's alpha for the two scales used in the questionnaire in which teachers were asked to rate the impact of participating in standardisation sessions and assessment training was .91 and .93 respectively, showing a high internal consistency. Spearman correlation analyses were also carried out to find relationships within the dataset.

Results

RQ1: Training in language assessment

Our first research question focused on the training in language assessment received by teachers in adult education in Spain, represented by the population in the two institutions surveyed. In order to answer this question, we looked at the percentage of teachers that had received training, the time when this training had taken place and the duration of the training in order to assess their level of assessment literacy. It was also relevant to find out about the organising institutions as well as whether the training had been organised nationally or had taken place abroad, in order to determine how training was being made accessible. Furthermore, it was considered important to examine the contents covered, to analyse the focus that had been given to the training and the way in which it had contributed to their assessment literacy.

Table 1. Percentage of teachers receiving training in language assessment

Total number of teachers that have received training	EOI teachers	ACLES teachers
76.5% (n=235)	75% (n=84)	78% (n=151)

As Table 1 shows, three-quarters of the teachers had received training in language assessment. A further question enquired about the period in which this training had taken place, and 74.9% of the teachers reported that they had attended training in the previous 1-3 years.

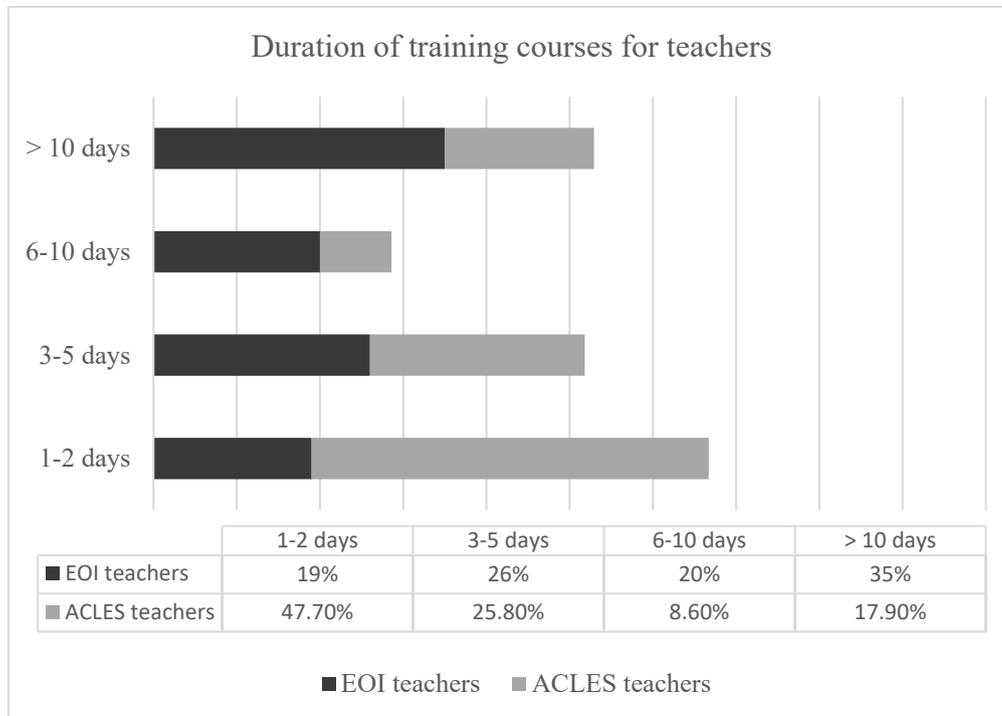


Figure 1. Duration of training received by teachers

As for the duration of the courses (Figure 1), they seemed to vary in length, with introductory courses as short as 1-2 days and longer courses covering content in more depth. As illustrated in the figure above, there is a higher prevalence of introductory courses among the teachers belonging to ACLES universities, while EOI teachers favoured longer courses.



Figure 2. Institutions organising the training received by teachers

As illustrated in Figure 2, training sessions were reported as largely organised by the institutions themselves or by their leading bodies, with some courses organised by other national organisations or by international organisations. Both in the case of EOIs and in the case of university language centres, teachers had also taken it upon themselves to organise training in their own teaching centres. The institutional influence in the encouragement of language assessment is evident in the large percentage of courses organised, as well as in their nature, as they were also identified as compulsory by some of the participants in the study. Furthermore, teachers gained access to courses organised at national level, with fewer teachers in both institutions having received training abroad (11% of EOI teachers and 8% of university teachers).

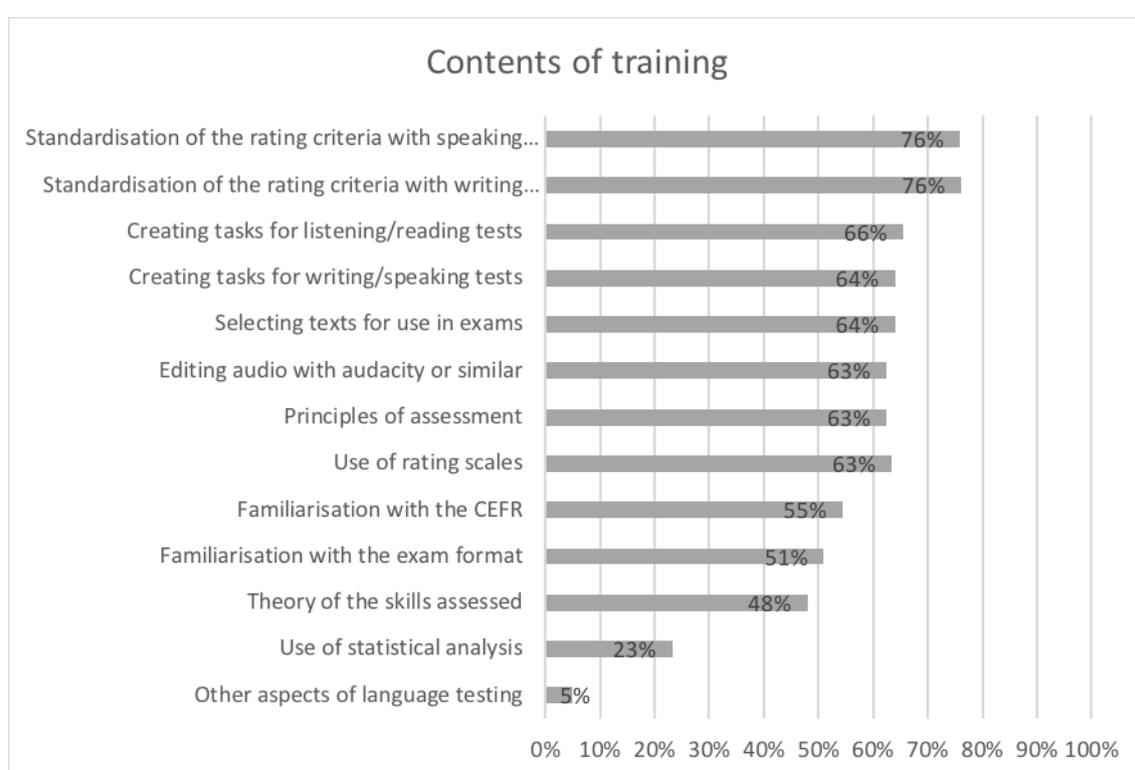


Figure 3. Contents of the training received

Table 2. Contents of the training received and teachers attending per institution

Contents covered	EOI teachers *	ACLES teachers
Creating tasks for listening/reading tests	72% (n=61)	59% (n=113)
Creating tasks for writing/speaking tests	65% (n=55)	63% (n=121)
Selecting texts for use in exams	64% (n=54)	64% (n=123)
Theory of the skills assessed	60% (n=51)	36% (n=70)
Editing audio with audacity or similar	65% (n=55)	60% (n=116)
Principles of assessment	85% (n=72)	40% (n=78)

Familiarisation with the CEFR	81% (n=69)	28% (n=54)
Standardisation of the rating criteria with speaking samples	94% (n=107)	57.8% (n=111)
Standardisation of the rating criteria with writing samples	94% (n=107)	58.3% (n=112)
Familiarisation with the exam format	75% (n=64)	27% (n=53)
Use of rating scales	81% (n=69)	45.8% (n=88)
Use of statistical analysis	0	46.4% (n=89)
Other aspects of language testing	6% (n=7)	3.6% (n=7)

Note. Not all survey respondents provided data for every question

The contents of the training received, illustrated by Figure 3 and Table 2, indicated an emphasis on the skill-based know-how that would allow teachers to take part in the development and critical assessment of standardised tests. However, principles and theory of testing that could help them use the exams developed fairly and in relation to their context were also included. The differences in content between EOI and ACLES illustrate the profile of the training organised by the two institutions, in the sense that the content has been suited to the professional needs of their teachers. Since the approach to the development and implementation of exams is different, the needs of the teachers differ and each institution designed courses to adapt to these needs. An example of this is familiarisation with the CEFR, which is not covered for teachers working in universities as it is already part of their training as teachers for the university, but is covered for EOI teachers as part of their standardisation sessions. Likewise, we see that in the case of standardisation of rating criteria, EOI teachers were involved to a larger degree than the ACLES teachers. This can be due to the fact that EOI teachers largely organise their own standardisation sessions, while ACLES follows a more centralised, top-down approach. Conversely, the use of statistical analysis is more prevalent for ACLES teachers, as it is a teacher responsibility, unlike for EOI teachers, whose analyses are centralised.

RQ2: Impact of training

Our second research question concerned the impact that the courses had had on the teachers' work. Impact was understood as positive impact and again the goal was to shed light on the teachers' perceptions about the field of language testing, which, according to the literature review above, tend to be that testing is a foreign field solely for specialists in the area. Our questions were intended to determine whether the amount and content of training could have an impact by providing them with experiences that were applicable to their work.

Table 3. Impact of training in work activities for EOI and ACLES teachers

Areas of work impacted by the training	Percentage of EOI teachers perceiving some or large impact in these areas	Percentage of university teachers perceiving some or large impact in these areas
Creating tasks for exams and assessing candidates' speaking and writing performances	83% (n=60)	83% (n=162)
Creating tasks for use in the classroom	75% (n=56)	85% (n=164)
General teaching practice	82% (n=65)	71% (n=137)

Note. Not all EOI survey respondents provided data for every question

Table 3 illustrates the percentage of teachers reporting a positive impact in their work. The first category impacted presents a compounded percentage of the teachers' impressions on the impact of the training on all activities related to developing and administering standardised tasks. The second category concerns the positive impact perceived on their general teaching practice, which comprises language teaching and other forms of assessment. All teachers concurred in identifying a clear effect of the training on the area of their work most directly related to the contents of the training: creating tasks for exams. However, the percentage of teachers who reported a positive effect in their general teaching practice was also large. Since we were interested in determining whether the length of the courses was related to the positive impact reported, Spearman correlation analyses were carried out and indicated that there were significant yet weak positive correlations between length of training and creating tasks for exams and assessing candidates' speaking and writing performances ($r_s(160)=.155$, $p < .001$) and length of training and positive impact in teaching practice, ($r_s(157)=.114$, $p < .001$). The weakness of the correlations leads us to conclude that length of courses was not a good indicator of positive impact on teachers' practices.

Our survey included an open question to further analyse teachers' perceptions. The questionnaire was anonymous and sent to a large number of participants, which leads us to believe that teachers' responses were honest and not swayed by external influences. Qualitative analyses of the answers indicated that the teachers' responses generally agreed that training had had a positive impact. They emphasised that training led to a self-reflection process based on the "experiences of other teachers", "insights gained in the factors that can affect exam candidates" and the importance of "obtaining a different perspective of what we do". In the case of teachers of languages other than English, they also emphasised the importance of "having the opportunity to work with teachers in your field", and "applying course content to different languages". Nevertheless, some of the comments were in line with the cautions in the literature about the foreignness of the field (Alderson, 2001; Davies, 2008; Fulcher, 2012; Pill &

Harding, 2013; Taylor, 2009), as some teachers made comments such as “fortunately, it has had no impact on my teaching practice” and that “time dedicated to exams is taken from actual teaching”.

RQ3: Further training needs

Our third research question concerned the self-perceived needs of teachers as regards training in assessment. The questionnaire asked them to identify the areas in which they needed training from a list of training course contents obtained from a literature review (Bailey & Brown, 1996; Brown & Bailey, 2008) and the contents of the courses already taking place in the institutions surveyed. The areas of interest for training and the percentage of teachers perceiving a need for further training in each area is illustrated in Figure 4.

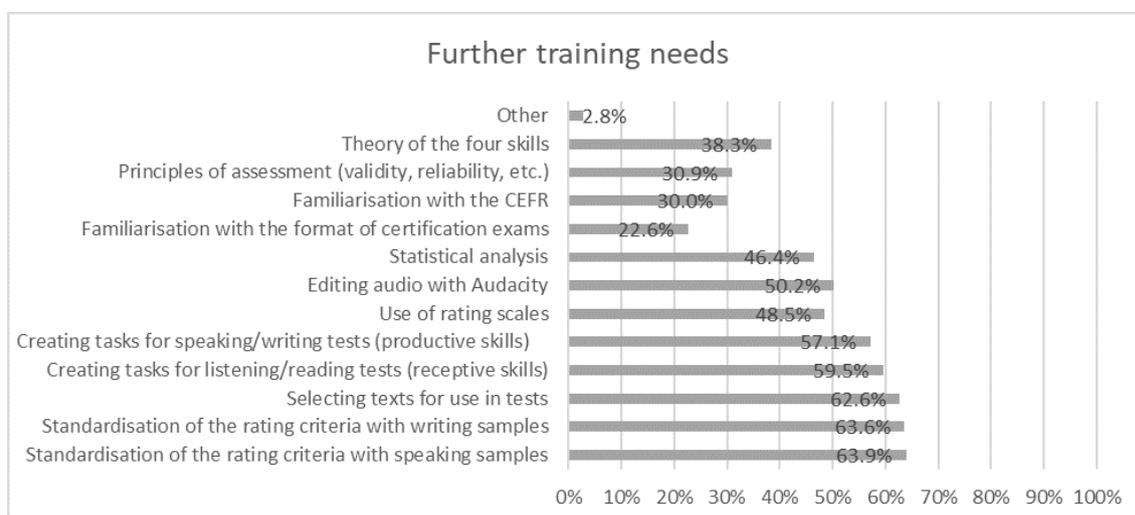


Figure 4. Further training needs

The teachers reported that they needed further training relating to how-to skills but also considered that the principles behind the skills were important, as seen by the percentages interested in the principles of assessment (30.9%) or the theory behind the assessment of the four skills (38.3%). In order to analyse whether the future needs perceived by the teachers were related to the training they had already received, we performed a correlation analysis that indicated that there was a significant and strong positive association between the specific contents of the training received and the participants' perception of needing further training in these specific contents ($r_s(252) = .884$, $p < .001$), that is, the more they learnt about specific topics, the more they indicated these topics as future training needs.

Since our question referred not only to the contents of the training but also to their needs regarding time constraints, a question had been included on the

amount of time they would be willing to dedicate to training activities. The results are illustrated in Figure 5.



Figure 5. Time available for training in language assessment

Most teachers opted for shorter courses with a maximum of 30 hours, although a percentage of them were willing to dedicate as much time as possible to training. Again, we were interested in whether the hours of training received and the positive impact on their work could be associated to the time they were willing to spend on further training. Spearman correlation analyses indicated that there was a significant but weak positive association between the training received and the time they considered they had available for future training ($r_s(253) = .267, p < .001$) and a negative, non-significant correlation between the impact of the training received and their availability for further training ($r_s(171) = -.012, p < .878$). The results of the correlation analyses pointed towards moderate to no association between the number of hours of training received and the impact of the training, as well as between the hours of training and the number of hours teachers could spend on future training. Lack of time was identified as a problem amongst teachers, which seemed to be supported by the preferred methods for training. These pointed towards a preference for tutored online courses or hybrid face-to-face and online courses (72.2%) over face-to-face courses (28.8%).

Summarising our results, we found that a large majority of teachers of adult language education in Spain had received specific training in language testing and standardisation of rating criteria in particular. The length varied according to the topic but training sessions of 1-2 days were most common, as they seemed to address specific topics and fit in their working schedules. Most training was

organised by the leading institutions and presented as compulsory for the teachers involved in standardised testing. The training sessions were part of institutional efforts to develop good practices in language testing and thus additional training efforts (sessions organised by teachers) were encouraged and facilitated. This was also corroborated by the timeframe of the sessions, as most training had occurred in the previous 1-3 years, coinciding with legislative changes that made standardised testing part of the teachers' responsibilities. The analysis of training contents indicated a prevalence of how-to skills, although broader aspects of LAL related to principles such as theory of the four skills and principles of assessment were not excluded. Furthermore, the differences between the training offered to teachers in the two institutions suggest that the institutions tailored the contents to the specific needs of their teachers. In addition to that, most training had occurred in Spain, suggesting context-bound approaches to the contents covered.

Our second question was related to the perceived positive impact that the training had had on the teachers' work. The results obtained indicated that a large percentage of the teachers (81.5%) perceived a positive effect on their work as test developers as well as on their general teaching practice (76.6%). The qualitative analysis of their comments identified some resistance to acquiring language assessment literacy on standardised tests but in general confirmed the quantitative data and hinted at the possibility of the training having an impact on classroom teaching and assessment. A correlation analysis between the length of the training and the impact in both areas mentioned showed a significant positive association but too weak to allow us to extract any conclusions in this respect.

Conversely, our third question identified the self-perceived training needs of teachers, which coincided in nature (favouring know-how skills but including broader principles) with the contents of the training received. A correlation analysis between the contents of the training received and self-perceived future needs indicated a strong positive correlation. Therefore, although the number of hours dedicated to training did not show a strong association with the impact on the teachers' work, the contents included in the training did encourage further interest amongst the participants in the survey. As for the time they were willing to dedicate to further training, the results were encouraging, with a majority of teachers willing to spend more than 20 hours. This willingness was shown to be moderately associated to the training already received but not to the impact perceived from such training. As for factors hampering training, lack of time seemed to be an issue, as indicated by the preference for flexible online or hybrid face-to-face and online methods.

Discussion

The aim of the present study is to contribute to the scholarship about language assessment literacy in teachers by investigating (a) the extent and profile of LAL training in teachers of adult education in Spain, (b) the impact of LAL training as perceived by the teachers, and (c) the perceived further needs as regards training in language testing.

Overall, our results indicated that language assessment literacy as regards standardised testing amongst language teachers in adult education in Spain had increased considerably in the three years prior to the study. In Spain, regional ministries of education, the labour market and the conference of university rectors have clearly influenced institutions to encourage LAL amongst their teachers and shaped the needs of teachers, who now perceive language assessment literacy as important for their professional practice. This period also coincided with the time in which governmental and educational regulations modified exam structures and required a larger number of teachers to take a more active role in test development within their institutions. The use of tests to improve educational outcomes is not new and has been reported by Spolsky (1995) and Brindley (2008) as a common practice of educational authorities. Furthermore, Malone (2008) and McNamara and Roever (2006) identified how tests were used for the purposes of accountability, both in the United States and in Europe. As a consequence of this, teachers have limited decision power on how tests are going to be used (Fulcher, 2012). The situation in Spain seems to mirror this tendency and a large number of teachers have been involved in training courses as part of an institutional effort to implement testing practices as educational policies and accountability measures. The fact that countries in which there is no clear policy concerning language assessment in adult education, as in the case of China, have comparatively lower levels of LAL amongst their teachers (Xu and Brown, 2017) could support our assumption that institutional involvement and the use of tests as policy tools act as a driving force for enhancing LAL in teachers.

As for the profile of the training, our results hinted at the strong influence of contextual factors, as shown by the prevalence of standardised testing and how-to skills in the training in Spain. This tendency for LAL efforts to be related to educational contexts and institutional involvement is mirrored in other countries. As an example, Vogt and Tsagari (2014), in their study on assessment literacy in seven European countries, identified that in Greece, the institutional context influenced teachers' LAL needs by requiring placement and achievement tests. On the other hand, in Germany, the school-leaving certificate's focus on written skills impacted on teachers' knowledge of assessment - or lack of thereof - and thus showed their preference for training in skills-based assessment. In the case

of Spain, the need for external language certificates and the institutional efforts to implement good practices in the development of such certificates placed the focus on proficiency examinations and on standardisation processes, which is reflected in the contents of the courses. Notwithstanding the different foci, our study rendered similar results to Hasselgreen et al. (2004), Guerin (2010), Vogt and Tsagari (2014), Brunfaut and Harding (2018) and Vogt et al. (2020) in that teachers expressed interest in training in many different areas of assessment based on their contexts, as well as increased support and tools, as found also by Baker and Riches (2018) in their study with Haitian teachers.

The impact of the training on their daily work was also of interest for the purpose of our study. Large efforts were dedicated to organising and financing training by both institutions, and the extent to which this training was perceived by teachers as something positive and thus beneficial for their work, or having an impact on their daily practice, had not been previously investigated. A large percentage of respondents believed that the training had an impact not only on assessment-related tasks, but also on their general teaching practice. This is in line with findings from similar studies in contexts in which training in assessment was also made available to teachers on a large scale, such as in the case of the Austrian Matura reform project (Konrad et al. 2018), which found that language assessment training for teachers led to a positive impact on classroom testing and a deeper understanding of the standard to be measured. Our results indicated that the teachers mostly appreciated how the training had helped widen their views by offering them a different perspective of their work. These findings seem to suggest that language assessment literacy could be moving towards becoming a field not limited to testing specialists. Length of training did not seem to affect the impact perceived, which could be attributable to the fact that most of the teachers were new to training in assessment and thus still discovering the applications of the field to their everyday work in the classroom.

As for the needs of teachers regarding LAL, our results again confirm the results of other studies, such as Deluca and Klinger's finding (2010) that teachers tend to express a need for general training in all aspects of language assessment. Again, there seems to be a general tendency towards a broader understanding of assessment and a gap for further training that has not yet been filled, or at least the results of the different surveys in the literature seem to indicate so. In the case of further training needs, context again plays a role, which can be seen in Spanish teachers' indication that further training is needed in subjects related to standardisation and in their interest in linking their exams with the CEFR, something that was also present in Brunfaut and Harding's (2018) study in Luxembourg. The reason seems to be very similar in both countries; institutions and the general population want to be able to use exams internationally, and linkage to a common framework and implementation of international good

practices are central to this purpose. Our questions regarding time available for training identified one of the challenges reviewed in the literature when dealing with training teachers in assessment, which is the lack of resources and time available. This suggests that in most cases, training is an additional burden to their workload. As mentioned previously, testing is still perceived as a separate field from teaching, and despite teachers identifying an impact on their work, not enough time seems to be allocated by the institutions for teachers to be able to develop professionally in this field. Future language courses could take this into consideration and limit the amount of teaching hours. Face-to-face training can be supplemented with self-study or online courses. Training in language assessment can be delivered not only through text-based materials, but also via other channels, such as the development of assessment instruments as described by Levi and Inbar-Lourie (2020) to help design LAL training that is process-oriented. Malone (2008) highlighted the contribution made by face-to-face workshops and self-instructional options, which often make use of new technologies and permit a wide range of materials to be made easily available (e.g., free downloading from the internet). An example of this is the Council of Europe's website materials associated with the Common European Framework of Reference for Languages (CEFR) (<https://www.coe.int/en/web/common-european-framework-reference-languages/resources>), as well as those provided by the Centre for Canadian Language Benchmarks (CLB) (<https://www.language.ca/resourceexpertise/>) or the materials offered by the Teachers' Assessment Literacy Enhancement (TALE) project (<https://taleproject.eu/>).

Conclusion

The results of this study reveal how training in assessment has reached a large number of teachers working in adult language education in Spain. Institutional involvement and changes in the legislative framework have acted as a catalyst for a large number of different stakeholders to become involved in high-stakes testing and to implement the measures needed to provide training for teachers. Language assessment is starting to be perceived as a wider field that can meld into the classroom and be beneficial in improving teaching practices by aligning these with learning and assessment. Furthermore, the contents of the training delivered and the needs expressed by the participants shed some light on the strong influence of contextual factors and on the way tests are used in each context. Assessment becomes a part of educational policy, and in a globalised economy with increased international mobility, transparency of testing practices and teacher involvement in these practices is important if the results of tests are to be used for achieving accountable and quality language education.

There are a number of limitations to this study that need to be acknowledged. Although the participants consisted of a sample population that was representative of professionals involved in assessment in both institutions, and their direct involvement in the survey contributed to the validity and reliability of the results, these were limited to the perceptions of the group of teachers surveyed. Although perceptions are subjective, they are important because they show how willing teachers may be to continue to participate in language assessment literacy training. Despite these limitations, our results could encourage further studies on the impact of test use by institutions and on the broadening of the scope of LAL.

References

- Alderson, J.C. (2001). Testing is too important to be left to the tester. Paper presented at the 3rd Annual Language Testing Symposium, Dubai, United Arab Emirates, <http://www.lancaster.ac.uk/fass/projects/examreform/Media/Article02.pdf>
- Alderson, J. C., Clapham, C., & Wall, D. (1995). *Language test construction and evaluation*. Cambridge University Press.
- Bailey, K. M., & Brown, J. D. (1996). Language testing courses: What are they? In A. Cumming & R. Berwick (Eds.), *Validation in language testing* (pp. 236–256). Multilingual Matters.
- 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. <https://doi.org/10.1177/0265532217716732>
- Brindley, G. (2008). Educational reform and language testing. In E. Shohamy & N. H. Hornberger (Eds.), *Encyclopedia of language and education: Vol. 7. Language testing and assessment* (2nd ed., pp. 365–378). Springer.
- Brown, G. T. (2011). Teachers' conceptions of assessment: Comparing primary and secondary teachers in New Zealand. *Assessment Matters*, 3, 45-70. <https://doi.org/10.18296/am.0097>
- Brown, J. D., & Bailey, K. M. (2008). Language testing courses: What are they in 2007? *Language Testing*, 25(3), 349-383. <https://doi.org/10.1177>
- Brunfaut, T., & Harding, L. (2018). Teachers setting the assessment (literacy) agenda: A case study of a teacher-led national test development project in Luxembourg. In D. Xerri & P. V. Briffa(Eds.), *Teacher involvement in high-stakes language testing* (pp. 155-172). Springer.

- Coombe, C., Vafadar, H., & Mohebbi, H. (2020). Language assessment literacy: What do we need to learn, unlearn, and relearn? *Language Testing in Asia*, 10(1). <https://doi.org/10.1186/s40468-020-00101-6>
- Council of Europe (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge University Press.
- Council of Europe (2020), *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume*. Council of Europe Publishing Available at www.coe.int/lang-cefr.
- Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing*, 25(3), 327-347. <https://doi.org/10.1177/0265532208090156>
- DeLuca, C., & Klinger, D. A. (2010). Assessment literacy development: Identifying gaps in teacher candidates' learning. *Assessment in Education: Principles, Policy & Practice*, 17(4), 419-438. <https://doi.org/10.1080/0969594x.2010.516643>
- 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. <https://doi.org/10.1177/0265532219826390>
- European Ministers of Education. (1999). *Bologna declaration*. <http://www.ehea.info/page-ministerial-conference-bologna-1999>
- Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9(2), 113-132. <https://doi.org/10.1080/15434303.2011.642041>
- Giraldo, F. (2020). A post-positivist and interpretive approach to researching teachers' language assessment literacy. *Profile: Issues in Teachers' Professional Development*, 22(1), 189-200. <https://doi.org/10.15446/profile.v22n1.78188>
- Guerin, E. M. (2010). Initial findings from a pilot Italian study of foreign language teachers' stated language assessment knowledge-base and needs. In G. Bota, H. Hargreaves, C. Lai, & R. Rong (Eds.), *Papers from the Lancaster University Postgraduate Conference in Linguistics & Language Teaching 2009: Volume 4*, (p. 80). Lancaster University. <http://wp.lancs.ac.uk/laelpgconference/papers-from-lael-pg/papers-from-lael-pg-2009-volume-4/>
- Hasselgreen, A., Carlsen, C., & Helness, H. (2004). *European survey of language testing and assessment needs. Part 1: General findings*. <http://www.ealta.eu.org/documents/resources/survey-report-pt1.pdf>
- Huhta, A., Hirvelä, T., & Banerjee, J. (2005). *European survey of language testing and assessment needs: Report: Part two—Regional findings*. http://users.jyu.fi/~huhta/ENLTA2/First_page.htm

- 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. <https://doi.org/10.1177/0265532209351431>
- Konrad, E., Spöttl, C., Holznecht, F., & Kremmel, B. (2018). The role of classroom teachers in standard setting and benchmarking. In D. Xerri & P. V. Briffa (Eds), *Teacher involvement in high-stakes language testing* (pp. 11-29). Springer.
- Lam, R. (2015). Language assessment training in Hong Kong: Implications for language assessment literacy. *Language Testing*, 32(2), 169-197. <https://doi.org/10.1177/0265532214554321>
- Levi, T., & Inbar-Lourie, O. (2020). Assessment literacy or language assessment literacy: Learning from the teachers. *Language Assessment Quarterly*, 17(2), 168–182. doi:10.1080/15434303.2019.1692347
- Malone, M. E. (2008). Training in language assessment. In E. Shohamy & N. H. Hornberger (Eds.), *Encyclopedia of language and education, Vol. 7: Language testing and assessment* (2nd ed., pp. 225–239). Springer.
- McNamara, T., & Roever, C. (2006). *Language testing: The social dimension*. John Wiley & Sons.
- Mertler, C. A. (2009). Teachers' assessment knowledge and their perceptions of the impact of classroom assessment professional development. *Improving Schools*, 12(2), 101-113. <https://doi.org/10.1177/1365480209105575>
- O'Loughlin, K. (2013). Developing the assessment literacy of university proficiency test users. *Language Testing*, 30(3), 363-380. <https://doi.org/10.1177/0265532213480336>
- Pill, J., & Harding, L. (2013). Defining the language assessment literacy gap: Evidence from a parliamentary inquiry. *Language Testing*, 30(3), 381-402. <https://doi.org/10.1177/0265532213480337>
- Popham, W. J. (2006). All about accountability / Needed: A dose of assessment literacy. *Educational Leadership*, 63(6), 84-85. <http://www.ascd.org/publications/educationalleadership/mar06/vol63/num06/Needed@-A-Dose-of-Assessment-Literacy.aspx>
- Rea-Dickins, P. (1997). So, why do we need relationships with stakeholders in language testing? A view from the UK. *Language Testing*, 14(3), 304-314. <https://doi.org/10.1177/026553229701400307>
- Rodriguez, C. (2015). *A study of the impact of language assessment literacy training on EOI teachers in Galicia, Spain* (Unpublished Master's thesis). Lancaster University.
- 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.
<https://doi.org/10.1177/0265532213480128>
- Spolsky, B. (1995). *Measured words: The development of objective language testing*. Oxford University Press
- Stabler-Havener, M. L. (2018). Defining, conceptualizing, problematizing, and assessing language teacher assessment literacy. *Studies in Applied Linguistics and TESOL*, 18(1), 1-22.
- Taylor, L. (2009). Developing assessment literacy. *Annual Review of Applied Linguistics*, 29, 21-36. <https://doi.org/10.1017/S0267190509090035>
- Taylor, L. (2013). Communicating the theory, practice and principles of language testing to test stakeholders: Some reflections. *Language Testing*, 30(3), 403-412. <https://doi.org/10.1177/0265532213480338>
- Vogt, K., & Tzagari, D. (2014). Assessment literacy of foreign language teachers: Findings of a European study. *Language Assessment Quarterly*, 11(4), 374-402. <https://doi.org/10.1080/15434303.2014.960046>
- Vogt, K., Tzagari, D. & Spanoudis, G. (2020) What do teachers think they want? A comparative study of in-service language teachers' beliefs on LAL training needs. *Language Assessment Quarterly*, 17(4), 386-409. <https://doi.org/10.1080/15434303.2020.1781128>
- Xu, Y., & Brown, G. T. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education*, 58, 149-162. <https://doi.org/10.1016/j.tate.2016.05.010>
- Xu, Y., & Brown, G. T. (2017). University English teacher assessment literacy: A survey-test report from China. *Papers in Language Testing and Assessment*, 6(1), 133-158.
- Yan, X., & Fan, J. (2020). "Am I qualified to be a language Tester?": Understanding the development of language assessment literacy across three stakeholder groups. *Language Testing*, 38(2), 219-246. <https://doi.org/10.1177/0265532220929924>

Appendix

Questionnaire about the assessment training of EOI teachers/ACLES Language Centre teachers

I would like to ask you to participate in this questionnaire about the training in language testing of teachers that work in EOIs in Galicia/ACLES language centres. This research aims to study the training needs in language testing of EOI/ACLES teachers and the impact that this training has on the work of teachers. This study will hopefully help to disseminate the important role that EOIs and Language Centres in universities have in language testing.

Your participation will allow teachers, schools and the administration to know our specific training needs in language testing, which can lead to training that is adapted to our context and designed to have the best possible impact on our work.

Your responses are absolutely confidential and anonymous. If you wish to receive information about the results of this study and you wouldn't mind being contacted to participate for an interview, you can choose to leave your contact details at the end of the survey. If you have any questions, please contact me.

1. Do you or have you worked in an EOI in Galicia/an ACLES Language Centre?
 - o Yes
 - o No

Training in language testing

2. How many times have you participated in a standardisation session of the rating criteria with teachers from your department or other departments?
 - o Never (*Go on to question 5*)
 - o 1-2 days
 - o 3-5 days
 - o 6-10 days
 - o More than 10 days
3. Mark from 0 to 3 the impact that these standardisation sessions have had on the following aspects of your work: (If you have not participated in an activity, please leave this line blank.)

	0 No impact	1 Small impact	2 Some impact	3 Large impact
Creating speaking tasks				
Creating writing tasks				
Selecting texts to use in exams				
Assessing the spoken performance of candidates in exams				
Assessing the written performance of candidates in exams				
Creating tasks for use in the classroom				
General teaching practice				

4. Would you like to add anything about the impact of standardisation sessions?

5. Have you received any type of training in language testing? (Training in language testing defined here as formal instruction in a course, workshop, etc. on the theory or practice of testing and/or familiarisation with the CEFR.)

- Yes
- No (*Go on to question 12*)

6. Number of days in which you participated in a language testing training event:

- 1-2 days
- 3-5 days
- 6-10 days

- o More than 10 days
7. In what period did this training occur? (Mark all of the correct options.)
- o In the last 1-3 years
 - o In the last 3-8 years
 - o More than 8 years ago
8. More information about these training events: (Please detail briefly where these training events occurred and who they were organised by)
-
9. Which of the following aspects were dealt with in this training? (Mark all of the options that were part of this training.)
- o Creating tasks for reading/listening tests (receptive skills)
 - o Creating tasks for writing/speaking tests (productive skills)
 - o Selecting texts for use in exams
 - o Theory of the skills assessed (reading, writing, listening, and speaking)
 - o Editing audio with Audacity or similar
 - o Principles of assessment (validity, reliability, etc.)
 - o Familiarisation with the CEFR
 - o Familiarisation with the certification exam test format
 - o Use of rating scales
 - o Other: _____

10. Mark from 0 to 3 the impact that this training in language testing has had on the following aspects of your work: (If you have not participated in an activity, please leave this line blank.)

	0 No impact	1 Small impact	2 Some impact	3 Large impact
Creating tasks for reading/listening tests (receptive skills)				
Creating tasks for writing/speaking tests (productive skills)				
Selecting texts to use in exams				

Editing audio with Audacity				
Assessing the spoken performance of candidates in exams				
Assessing the written performance of candidates in exams				
Creating tasks for use in the classroom				
General teaching practice				

11. Would you like to add anything about the impact of this training in language testing?

Training needs in language testing

12. In which of the following aspects do you think you need more training? (Mark all of the options you consider you need more training in.)

- Creating tasks for reading/listening tests (receptive skills)
- Creating tasks for writing/speaking tests (productive skills)
- Selecting texts for use in exams
- Theory of the skills assessed (reading, writing, listening, and speaking)
- Editing audio with Audacity
- Principles of assessment (validity, reliability, etc.)
- Familiarisation with the CEFR
- Standardisation of the rating criteria with speaking samples
- Standardisation of the rating criteria with writing samples
- Familiarisation with the certification exam test format
- Use of rating scales
- Other: _____

13. How much time would you be willing to invest in training in language testing?

- 3-4 hours
- 5-10 hours

- o 11-20 hours
- o 20-30 hours
- o Other: _____

14. Which would be the most ideal format for this training?

- o Online course with a tutor
- o Online materials for self-study
- o Face-to-face course
- o Combination of online and face to face course
- o Other: _____

15. Would you like to add anything about your training needs in language testing?

Other Information

16. What EOI/Language Centre do you currently work at or have you last worked at?

17. Do you work in the main EOI or in an EOI section?

- o Main EOI
- o EOI section in the same city as the EOI
- o EOI section in a town different from main EOI
- o Not applicable

18. What language(s) do you teach?

19. How long have you taught at a/n EOI/Language Centre?

20. Have you taught in a secondary school? If so, how long have you taught in a secondary school?

21. Have you taught in a different sector? If so, how long have you taught in a different sector?

22. Which of the following activities have you done as part of your work at a/n EOI//Language Centre? (Mark all of the correct options.)

- o Creating tasks for standardised exams
- o Creating tasks for other exams
- o Administering standardised exams
- o Rating the spoken performance of candidates on standardised exams
- o Rating the written performance of candidates on standardised exams

23. Level of studies:

24. Gender:

25. Age:

26. If you wish to leave your contact information you can do so. You will receive further information about the results of this study and you may be contacted to participate in a follow-up interview. Your personal information will remain anonymous.

Thank you for your participation!