

**An Investigation of English as a Foreign Language Teachers' Assessment Literacy in  
Kazakhstan**

Dulat Irzhanov

Submitted in partial fulfilment of the requirements for the degree of

Master of Arts

in

Multilingual Education

Nazarbayev University Graduate School of Education

May, 2022

Word count: 16,982

### Author Agreement

By signing and submitting this license, I Duldat Irzhanov (the author or copyright owner) grant to Nazarbayev University (NU) the non-exclusive right to reproduce, convert (as defined below), and/or distribute my submission (including the abstract) worldwide in print and electronic format and in any medium, including but not limited to audio or video.

I agree that NU may, without changing the content, convert the submission to any medium or format for the purpose of preservation.

I also agree that NU may keep more than one copy of this submission for purposes of security, back-up and preservation.

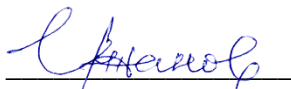
I confirm that the submission is my original work, and that I have the right to grant the rights contained in this license. I also confirm that my submission does not, to the best of my knowledge, infringe upon anyone's copyright.

If the submission contains material for which I do not hold copyright, I confirm that I have obtained the unrestricted permission of the copyright owner to grant NU the rights required by this license, and that such third-party owned material is clearly identified and acknowledged within the text or content of the submission.

IF THE SUBMISSION IS BASED UPON WORK THAT HAS BEEN SPONSORED OR SUPPORTED BY AN AGENCY OR ORGANIZATION OTHER THAN NU, I CONFIRM THAT I HAVE FULFILLED ANY RIGHT OF REVIEW OR OTHER OBLIGATIONS REQUIRED BY SUCH CONTRACT OR AGREEMENT.

NU will clearly identify my name(s) as the author(s) or owner(s) of the submission, and will not make any alteration, other than as allowed by this license, to your submission.

I hereby accept the terms of the above Author Agreement.

  
Author's signature:

27.05.2022  
Date:

**Declaration**

I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been submitted for the award of any other course or degree at NU or any other educational institution, except where due acknowledgement is made in the thesis. This thesis is the result of my own independent work, except where otherwise stated, and the views expressed here are my own.

Signed: 

Date: 27.05.2022

## Ethical Approval



53 Kabanbay Batyr Ave.  
010000 Astana,  
Republic of Kazakhstan  
15<sup>th</sup> October 2021

Dear Dulat Irzhanov

This letter now confirms that your research project entitled: **An investigation of English Foreign Language (EFL) teachers' assessment literacy in Kazakhstan** has been approved by the Graduate School of Education Ethics Committee of Nazarbayev University.

You may proceed with contacting your preferred research site and commencing your participant recruitment strategy.

Yours sincerely



Michelle Bedeker

On behalf of Zumrad Kataeva  
Chair of the GSE Ethics Committee  
Assistant Professor  
Graduate School of Education  
Nazarbayev University

Block C3, Room 5006  
Office: +7 (7172) 70 9371  
Mobile: +7 777 1929961  
email: [zumrad.kataeva@nu.edu.kz](mailto:zumrad.kataeva@nu.edu.kz)

## CITI Training Certificate



Completion Date 04-Sep-2021  
Expiration Date 03-Sep-2025  
Record ID 44755325

This is to certify that:

**Dulat Irzhanov**

Has completed the following CITI Program course:

Not valid for renewal of certification through CME.

**Responsible Research Training**

(Curriculum Group)

**Social, Behav, Edu, Etc**

(Course Learner Group)

**1 - Basic Course**

(Stage)

Under requirements set by:

**Nazarbayev University**

**CITI**  
Collaborative Institutional Training Initiative

Verify at [www.citiprogram.org/verify/?w1a8c07bc-6e4b-479f-8d5d-d7ecf25992ec-44755325](http://www.citiprogram.org/verify/?w1a8c07bc-6e4b-479f-8d5d-d7ecf25992ec-44755325)

### **Acknowledgements**

It is a genuine pleasure to express my deep sense of thanks and gratitude to my beloved thesis supervisor Dr. Michelle Bedeker. The completion of this study could not have been possible without her guidance and advice in all the stages of writing my project. I would also like to acknowledge and give my warmest thanks to professors of the Multilingual Education program at Nazarbayev University Graduate School of Education, who introduced the world of multilingualism, academia, and education to us.

Furthermore, I would also like to acknowledge the crucial role of my colleague Gulnara Duskaziyeva who made my academic journey become a reality. Several years ago, I could not even imagine that I, an ordinary English teacher (not even enough fluent in English), could manage such a study, and she is the one who motivated me to apply to this program.

Also, many thanks to my dear groupmates for unwavering support day in and day out during these unbelievable two years on Zoom and on campus, and for reminding me to have fun when I have been stressed out.

Finally, I am thankful to my family, without whom none of this would be possible. Thank you for believing in me and supporting me all the time.

### **Abstract**

In Kazakhstan, English as a foreign language (EFL) teachers seem to be in a challenging position because of the recent changes in the country's educational system, which requires teachers to radically change their values and beliefs about good teaching and assessment. In addition, although there is a growing demand to improve the quality of English teaching at schools, limited classroom-based research studies in Kazakhstan have focused on English as a foreign language (EFL) teachers' language assessment literacy (LAL), as well as the relationship between state-funded professional development programmes and teachers' LAL development is underexplored. It suggests that EFL teachers may lack the necessary knowledge, skills, and confidence when appropriating new assessment strategies in their classroom practices, which can be critical for implementing the envisaged educational changes successfully. Consequently, this research investigates how EFL teachers' LAL reflects the underpinnings of language assessment reform in the Kazakh context and aims to shed light on how state-funded assessment PD courses impact EFL teachers' LAL.

In this convergent parallel mixed-methods study, the researcher specifically examined the (a) content of language assessment PD training in Kazakhstan, (b) teachers' skills and knowledge about language assessment, as well as (c) their beliefs and attitudes about language assessment; in relation to Giraldo's (2018) descriptions for LAL dimensions, which were employed as the conceptual framework of this study. The study sample consisted of 119 participants in a teacher survey, three participants in teacher interviews, and three EFL teacher training course descriptions. Findings showed that state-funded EFL teacher PD courses (a) primarily focus on sharing assessment knowledge related to the recent school curriculum update (criteria-based assessment, formative and summative assessment), whereas (b) the significant component of LAL, the knowledge that includes applied linguistics such as

theories of second language acquisition, critical language testing, and multilingualism, is not present. It suggests that completing the courses does not sufficiently develop the teachers' LAL. The results echoed in both document analysis, teacher survey and teacher interviews. Finally, the study results indicate the needs of in-service EFL teachers that are not covered in the state PD curriculum and propose various implications and recommendations for developing teacher LAL in Kazakhstan.

*Keywords:* assessment literacy, language assessment literacy, EFL teachers' LAL level, training and development, reform and education

### Аңдатпа

Қазақстандағы білім беру жүйесіндегі мұғалімдерден өздерінің құндылықтары мен дұрыс оқыту мен бағалау туралы сенімдерін түбегейлі өзгертуді талап ететін соңғы реформалары аясында, елімізде ағылшын тілі (EFL) мұғалімдері қиын жағдайда болуы мүмкін екені белгілі. Сонымен қатар, мектептерде ағылшын тілін оқыту сапасын арттыруға сұраныстың артуына қарамастан, Қазақстанда EFL мұғалімдерінің тілдік бағалау құзыреттілігі (LAL) бойынша зерттеулер аз. Сондай-ақ, мұғалімдердің кәсіби біліктілігін арттырудағы мемлекеттік бағдарламалары мен EFL мұғалімдерінің LAL деңгейінің дамуы арасындағы байланыс толық зерттелмеген. Осының барлығы EFL мұғалімдерінде жаңа бағалау стратегияларын сынып тәжірибесінде қолдану үшін қажетті білімнің, дағдылардың және өзіне деген сенімділіктің болмауы мүмкін екеніне нұсқап, білім беру саласындағы көзделген өзгерістерді сәтті жүзеге асыруда маңызды рөл атқаруы мүмкін екендігін көрсетеді. Сондықтан, зерттеудің мақсаты – Қазақстанда EFL мұғалімдерінің LAL деңгейі тілді бағалау реформасының негіздерін қалай көрсететінін және мемлекеттік педагог біліктілігін арттыру курстары мұғалімдердің LAL деңгейіне қалай әсер ететінін анықтау болды.

Аталған конвергентті параллель аралас әдісті ғылыми жұмыс (а) Қазақстандағы тілді бағалау бойынша EFL мұғалімдеріне арналған біліктілікті арттыру курстарының мазмұнын, (ә) мұғалімдердің тілді бағалаудағы дағдылары мен білімдерін және (б) олардың тілді бағалау туралы сенімдері мен көзқарастарын зерттеді; зерттеудің тұжырымдамалық негізі ретінде Giraldo-ның (2019) LAL параметрлері сипаттамалары пайдаланылды. Зерттеу үлгісі 119 мұғалім сауалнамасы қатысушыларынан, үш мұғалім сұхбатына қатысушысынан және үш EFL мұғалімдеріне арналған мемлекеттік кәсіби біліктігін арттыру курстарының сипаттамасынан тұрды. Зерттеу нәтижелері EFL

мұғалімдеріне арналған мемлекеттік кәсіби біліктілігін арттыру курстары (а) ең алдымен мектептегі білім саласының соңғы өзгерістеріне (критериалды бағалау, формативті және жиынтық бағалау) қатысты бағалау туралы біліммен бөлісуге бағытталғанын және (ә) LAL-дың маңызды компоненті, екінші тілді меңгеру теориялары (SLA), тілді сыни тестілеу (CLT) және көптілділік (Multilingualism) сияқты қолданбалы лингвистика ұғымдарын қамтымайтынын көрсетті. Сәйкесінше, аталған курстардың мұғалімдер LAL-ның жеткілікті түрде дамуына әсер етпейтіні анықталды. Алынған нәтижелер құжаттарды талдауда да, мұғалімдермен жүргізілген сауалнама мен сұхбаттарда да көрініс тапты. Қорытындылай келе, зерттеу нәтижелері қазіргі EFL мұғалімдерінің мемлекеттік педагогтар кәсіби біліктілігін арттыру жоспарында қамтылмаған қажеттіліктерін көрсетіп, Қазақстандағы мұғалімдердің LAL деңгейін дамытудағы салдар мен орындалуы қажет әрекеттерді ұсынады.

### Аннотация

В свете недавних изменений в системе образования в Казахстане, которые требуют от учителей радикального изменения своих ценностей и убеждений в отношении хорошего преподавания и оценивания, учителя английского языка (EFL) в стране, возможно, находятся в сложном положении. Кроме того, несмотря на растущий спрос на повышение качества преподавания английского языка в школах, исследований, направленных на изучение компетенции в области оценивания языковых знаний (LAL) среди учителей EFL в Казахстане мало. Также, взаимосвязь между финансируемыми государством программами профессионального развития педагогов и развитием LAL учителей EFL изучена недостаточно. Все это указывает на то, что учителям EFL может не хватать необходимых знаний, навыков и уверенности в себе при применении новых стратегий оценки в своей классной практике, что может иметь решающее значение для успешного внедрения предусмотренных образовательных изменений. Следовательно, целью исследования выступает изучение того, как LAL учителей EFL отражает основы реформы языковой оценки в казахстанском контексте, и как финансируемые государством курсы повышения квалификации педагогов влияют на LAL учителей.

В рамках данного конвергентного параллельного исследования смешанного метода были изучены (а) содержание обучения профессионального развития педагогов по языковой оценке в Казахстане, (б) навыки и знания действующих учителей в области языковой оценки, а также (в) их убеждения и отношение к языку; в качестве концептуальной основы этого исследования использовались описания параметров LAL предложенным Giraldo (2018). Выборка исследования состояла из 119 участников опроса учителей, трех участников интервью с учителями и трех описаний курсов подготовки для учителей EFL. Результаты показали, что финансируемые государством

курсы повышения квалификации педагогов (а) в первую очередь направлены на предоставление знаний об оценке, связанной с недавним обновлением школьной программы (критериальное оценивание, формативная и суммативная оценка), (б) и в курсах отсутствует важный компонент LAL, который включает знания прикладной лингвистики, такие как теории овладения вторым языком, критическое языковое тестирование и полиязычие. Это позволяет предположить, что завершение курсов недостаточно развивает LAL учителей. Полученные результаты нашли отражение как в анализе документов, опросе учителей, так и в интервью с учителями. Наконец, результаты исследования указывают на потребности действующих учителей английского языка, которые не охвачены в государственной учебной программе повышения квалификации учителей, и предлагают различные последствия и рекомендации для развития LAL у учителей в Казахстане.

**Table of Contents**

Author Agreement .....	ii
Declaration.....	iii
Ethical Approval.....	iv
CITI Training Certificate.....	v
Acknowledgements .....	vi
Abstract.....	vii
List of Figures.....	xvii
List of Tables .....	xviii
Chapter 1: Introduction.....	1
The Problem Statement .....	3
Purpose Statement .....	3
Research Question.....	4
Significance of the Study .....	4
Outline of the Study .....	5
Definition of Terms.....	5
Conclusion.....	6
Chapter 2: Literature Review .....	7
Introduction to Assessment in Kazakhstan .....	7
Developments within the field of language testing .....	8
Critical language testing debates.....	9
The multilingual turn and testing debates .....	11
The emergence of language assessment literacy (LAL).....	12
Empirical studies about AL and LAL .....	13

Summary .....	16
Chapter 3: Methodology .....	17
Research design and method .....	17
Sampling and Research Instruments .....	18
Data Collection Procedures .....	20
Data Analysis Procedure .....	21
Ethical Considerations.....	22
Confidentiality and Anonymity.....	23
Risks and Benefits.....	24
Summary .....	24
Chapter 4: Findings .....	25
Document Analysis .....	25
Findings of the Website Analysis .....	26
Findings from Course Content .....	28
Summary of findings.....	30
Quantitative Analysis .....	31
Presentation of Survey Findings .....	31
Demographics .....	31
Quantitative Analysis of Likert-type items .....	33
Presentation of Qualitative Data.....	45
Presentation of the Interview Results.....	45
Participants' Profile.....	45
Teachers' self-perceived assessment knowledge .....	46
Managing assessment in practice .....	48
Professional Development and EFL Teachers' Professional Identity.....	49

Developing teachers' language assessment skills for classroom-based practices.....	51
Summary of interview findings.....	52
Summary of findings across the three data sets .....	53
Conclusion.....	58
Chapter 5: Discussion.....	59
Research Question 1: What is the content of language assessment PD training in Kazakhstan?.....	59
Research Question 2: What are teachers' skills and knowledge about language assessment? .....	61
Research Question 3: What are teachers' beliefs and attitudes about language assessment?..	62
Main Research Question: To what extent do teachers' language assessment literacy (LAL) reflect the underpinnings of language teaching reform? .....	64
Conclusion.....	65
Chapter 6: Conclusion .....	67
Main Conclusions of the Study .....	68
The content of language assessment PD training in Kazakhstan.....	68
Teachers' skills and knowledge about language assessment.....	68
The relationship between teachers' language assessment literacy (LAL) and language teaching reform .....	69
Limitations of the Study .....	70
Implications and Recommendations .....	71
Practice.....	71
Future Research.....	71

<b>References</b> .....	73
Appendix A.....	85
Appendix B.....	88
Appendix C.....	89
Appendix D.....	91
Appendix E.....	100

## List of Figures

Figure 1.....	18
Figure 2.....	29
Figure 3.....	38
Figure 4.....	40

## List of Tables

Table 1 .....	21
Table 2 .....	26
Table 3 .....	27
Table 4 .....	32
Table 5 .....	33
Table 6 .....	34
Table 7 .....	36
Table 8 .....	42
Table 9 .....	42
Table 10 .....	43
Table 11 .....	43
Table 12 .....	46
Table 13 .....	55

## Chapter 1: Introduction

One of the crucial roles of English as a Foreign Language (EFL) teachers is assessing students' performance because it plays a critical role in improving learning (Black & William, 2010). Moreover, language tests today significantly influence the educational system, especially teachers and students (Fulcher, 2012). Therefore, language teachers and other stakeholders who make decisions based on assessment data should be good at interpreting, designing, implementing, and evaluating language assessment and analysing the implications of their assessment-based actions (Scarino, 2013). Consequently, English as a foreign language (EFL) teachers should be skilled in multiple aspects of assessment to achieve their intended teaching goals. However, most studies on EFL teachers' assessment strategies found that teachers were unprepared to implement classroom assessments (Alkharusi, 2011; Vogt & Tsagari, 2014; Xu & Brown, 2017). For this reason, researchers call for further studies targeting language teachers' assessment practices, knowledge, and needs.

Stiggins (1991) coined the term assessment literacy (AL) to shed light on teachers' assessment skills, knowledge, and practices. For Stiggins, the term AL encapsulates how teachers differentiate between assessment goals, collect valid assessment data, and comprehend how to employ assessment techniques and data to facilitate students' development and awareness of their educational progress. Interestingly, Stiggins argued that only assessment literate educators would understand how to produce and use high-quality assessment data, thus contributing to "important instructional and policy decisions" (Stiggins, 1991, p. 535).

More than a decade later, Popham (2004) argued that assessment training to develop assessment literate educators was still lacking, which was tantamount to "professional suicide" (p. 82). The situation improved during the following decade since many researchers,

governments, language testers, and testing agencies advocated measuring and improving teachers' AL, most notably language assessment literacy (LAL) (Fulcher, 2012; Inbar-Lourie et al., 2013; Lam, 2015; Malone, 2013; Taylor, 2013). However, recent research studies highlight that most language teachers are still not ready to use and interpret assessment data to enhance their teaching and instruction (Lam, 2015; Vogt & Tsagari, 2014). Globally, LAL is a critical component of education systems because language is the tool to develop assessments, it is the object of assessment, and it is the mediator of content in other subjects (Alkharusi et al., 2012; Bachman & Palmer, 2010; Davies, 2008; Fulcher, 2012; Inbar-Lourie et al., 2013; Malone, 2013; Scarino, 2013; Taylor, 2009; 2013; Vogt & Tsagari, 2014; Xu & Brown, 2017). As a result, LAL has "tremendous potential to enhance student learning and change a whole assessment culture" (Hidri, 2020, p. xiv). Given the educational reform in the Kazakhstani context and the significant 'retraining' of EFL teachers, it is imperative to shed light on EFL teachers' LAL.

In 2016, the Ministry of Education and Science of Kazakhstan (MoES, 2015) launched significant educational reforms to update the language-in-education policy, curriculum, and assessment methods for school education. However, the educational reforms shifted from previous behaviourist teaching and testing methods, requiring teachers to radically change their values and beliefs about good teaching and assessment (Karabassova, 2020). To support and develop in-service teachers for this rapid change in teaching and assessment methodology, the government has provided school administrators and teachers with state-funded professional development courses, workshops, and seminars, combined with attractive financial bonuses (MoES, 2015). In addition, teacher preparation programs at Kazakhstani universities have had to make the necessary changes to align their programs with these educational reforms (MoES,

2015). Therefore, LAL research is timely because EFL teachers' assessment skills can be critical for implementing the envisaged educational changes successfully.

### **The Problem Statement**

Kazakhstani educational reform transformed the assessment system from a previous norm-referenced underpinning that compared students' results against the achievements of other students to a formative assessment system (Karabassova, 2021; Shamshidinova et al., 2014). However, limited classroom-based research studies in Kazakhstan have focused on EFL teachers' language assessment literacy (LAL); more importantly, the relationship between state-funded professional development programmes and teachers' LAL development is underexplored. EFL teachers' language assessment literacy is essential because shifting from testing decontextualised grammar rules to formative assessment could be challenging for many language teachers. For this reason, the Kazakh educational reform initiatives were critiqued as a "half-baked policy" because they resulted in teachers' misinterpreting key terms such as formative assessment, summative assessment, and differentiated teaching (Karabassova, 2020, p. 37). Consequently, it is essential to explore Kazakhstani EFL teachers' understanding of how new assessment models, such as criteria-based assessments, could unintentionally contribute to their EFL implementation challenges.

### **Purpose Statement**

Undoubtedly, there is a growing interest in EFL teachers in the Kazakh context as the key figures in implementing the trilingual education policy (MoES, 2015). The reason for this is that English has taken on a prominent role as a subject and a medium of instruction in STEM subjects. Therefore, Kazakhstan has developed a professional development system to support teachers' professional growth: the JSC National Centre for Professional Development "Orleu"

and the Centre of Excellence (CoE) of the Autonomous Educational Organisation Nazarbayev Intellectual Schools; both with branches in all regions of the country. However, recent studies have highlighted that PD courses have not provided EFL teachers with the practical classroom tools and methods to implement formative, summative, and criteria-based assessments (Akhmedina, 2017; Bondarenko, 2019). These studies suggest that EFL teachers may lack the necessary knowledge, skills, and confidence when appropriating new assessment strategies in their classroom practices. Consequently, this research investigates how EFL teachers' LAL reflects the underpinnings of language assessment reform in the Kazakh context. For this reason, the purpose of this study is to shed light on how state-funded assessment PD courses impact EFL teachers' language assessment literacy.

### **Research Question**

**The central research question of this study is:**

To what extent do teachers' language assessment literacy (LAL) reflect the underpinnings of language assessment reform?

**The following sub-questions further guide it:**

1. What is the content of language assessment PD training in Kazakhstan?
2. What are teachers' skills and knowledge about language assessment?
3. What are teachers' beliefs and attitudes about language assessment?

### **Significance of the Study**

Language assessment, specifically EFL teachers' language assessment literacy, is highly relevant in the current Kazakhstani educational context because it can improve the quality of teaching and language assessment. For this reason, the study's results would be beneficial to all stakeholders in the education system, including learners, teachers, school administrators, and

even policymakers. First, the EFL teachers might better understand their current language assessment literacy (LAL) associated with its core components, namely knowledge, skills, and principles (Davies, 2008; Giraldo, 2018). Secondly, it may reveal to policymakers and developers of the teacher professional development curriculum what EFL teachers' assessment needs are, resulting in potential updates in such programs. Finally, it will be important that emerging graduate students or researchers read this study because it can be helpful when planning further research, maybe diachronic studies focusing on the evaluation of assessment-specific PD courses' for EFL teachers.

### **Outline of the Study**

The thesis consists of six chapters. First, Chapter 1 introduced and provided the background about language assessment literacy and curriculum reform in Kazakhstan and the study's purpose and significance. Chapter 2 will review the literature about developments within language testing. The third chapter focuses on the research design, methodology, and ethical considerations that guided this research study. Then Chapter 4 presents the data and findings of the research. Finally, Chapter 5 discusses the results, and Chapter 6 concludes the study with a summary of the main conclusions and the study limitations and makes recommendations for further research.

### **Definition of Terms**

The following definitions were used for this study:

*Assessment Literacy (AL)* is a set of beliefs, knowledge, and practices about assessment (Stiggins, 1991).

*Language Assessment Literacy (LAL)* is LAL is different from assessment literacy because the assessment of language knowledge and communicative competence has its unique complexities (Harding & Kremmel, 2016; Jeong, 2013).

*Formative Assessment* is a continuous process during learning and teaching in which teachers use assessment-elicited evidence of students' learning to adjust their pedagogy to encourage students learning (Popham, 2008, p. 6).

*Summative Assessment* is an end of a course evaluation to judge if learning goals were achieved, track improvement, certify achievement, and openly report achievement data (DuFour et al., 2010; Harlen, 2012; McMillan & Nash, 2000; Moss, 2013).

## **Conclusion**

This chapter has provided the study's background information, rationale, and significance. The next chapter will provide a critical review of the literature associated with assessment and language assessment literacy.

## **Chapter 2: Literature Review**

The research aims to identify English as a foreign language (EFL) teachers' language assessment literacy (LAL) and determine its correlation to the state-funded teacher training programs' aims and goals. The previous chapter provided a background of the study; this chapter will focus on the literature that underpinned my research project. I will first introduce the history of language testing in Kazakhstan (2.1) and then present information on how the field of language testing has developed with a specific focus on critical language testing debates (2.2). Furthermore, Section 2.3 in this chapter critically reviews the emergence of language assessment literacy and its key debates. This chapter concludes with empirical research studies about language assessment literacy in the USA, South America, Western Europe, and the Asian contexts.

### **Introduction to Assessment in Kazakhstan**

Before independence, Kazakhstan's education system was underpinned by Soviet values about teaching and assessment. For example, teachers' training and classroom practices were rooted in a traditional assessment system that valued regurgitation and memorisation of facts as evidence of successful academic learning (Bondarenko, 2019). The OECD (2014) stated that Kazakhstani teachers were still using a 1 to 5 scale for assessing students' learning achievements and "a set of general didactic criteria that were first developed during Soviet times" (p. 117). For this reason, the education system inherited from the Soviet era was criticised as too rigid and not relevant in an era signalling modernisation, global competitiveness, and free-market values (Tastanbekova, 2020). As a result, educational reform after independence signalled a dramatic break from soviet values requiring teachers to make a paradigm shift associated with their beliefs about language pedagogy and assessment.

To address the policy mishaps, the government has initiated two state-funded teacher training organisations asked to focus on the development of training programs to address the implementation gaps associated with curriculum reform and assessment. According to MoES (2020), since 2016, a total number of 317 271 educators have been participating in training, workshops, and seminars at “Orleu” (Orleu) and the Centre of Excellence (CoE) of the autonomous educational organisation Nazarbayev Intellectual Schools (p. 187). Orleu and CoE are state-funded organisations, and they lead the process of updating teachers' skills in Kazakhstan. For this reason, an investigation into their EFL assessment courses can illustrate how they contribute to transforming and developing teachers' LAL from previous traditional testing practices. In the Kazakhstani context investigating EFL teachers' LAL is critical because they are non-native English speakers. The extent they have been trained in multilingual theories and applied linguistics to implement new pedagogical and assessment approaches is questionable (Davies, 2008; Fulcher, 2012; Giraldo, 2018).

### **Developments within the field of language testing**

Educational testing has a long history which emerged during the 1500 century in Imperial Examination in China. Interestingly, this was the first centralised state-wide educational testing in history, at about the same time in northern Italy when schools started using the medieval Treviso test, the first test focusing on the educational process (O'Sullivan, 2002; Spolsky, 2017). Later in the 16th century, language assessment needs arose in different contexts such as Europe and the U.S. Its focus was on literacy testing of skills such as compositions, translations, and multiple-choice grammar and comprehension (Spolsky, 2017).

During the 19<sup>th</sup>-century, testing practices encouraged memorisation and regurgitation of grammar and decontextualised language knowledge which became pervasive because it was seen

as contributing to upward social mobility for learners from low socio-economic backgrounds (Latham, 1877; as cited in Spolsky, 2017, p. 381). However, large-scale testing requirements resulted in standardisation challenges, test validity, and fairness, which were of interest across three language testing phases: pre-scientific, psychometric-structuralist, and psycholinguistic-sociolinguistic language testing (Spolsky, 1977). The subjective nature of language tests characterised the first phase, where teachers were the primary assessors; they ignored test reliability and validity issues. Then, the psychometric-structuralist phase focused on objectivity in language assessment which piqued the interest of theoretical and applied linguists; still one of the primary issues in language testing (McNamara & Shohamy, 2008; Shohamy, 2001; Spolsky, 2007). Finally, the psycholinguistic-sociolinguistic phase concentrated on the test taker, the test designers, test purposes, and fairness in language testing; this phase highlighted that tests are powerful tools that hold academic, identity, and social justice implications (McNamara & Roever, 2006; Shohamy, 2001; Spolsky, 2008).

Since then, the social dimensions of language testing have sparked interest in diverse areas such as language testing washback, the effect of English language proficiency on students' performances, and recently the potential of multilingual assessment (Davies, 2004; 2008; Giraldo, 2018; Scarino, 2010; 2013; Shohamy, 2001). Consequently, the field of critical language testing emerged, pointing out that tests are not neutral, that the use of test results needs critical interrogation about who is assessed, in what language, and whose agendas or educational ideologies are advanced (Kramersch, 1993; 2014; Pennycook, 2001; Shohamy, 2017).

### **Critical language testing debates**

In the 1980s, the shift from decontextualised grammar testing resulted in a focus on learners' communicative competencies in language testing (Kramersch, 1993; Spolsky, 1977).

During this time, large-scale English language testing companies, such as the American Test of English as a Foreign Language (TOEFL) and the University of Cambridge Local Examination Syndicate (UCLES), emerged, interested in two goals. First, by making a profit from the growing demand for English language teaching and testing globally, and second by researching how to improve the quality of their tests (Bachman, 1990; Bachman & Palmer, 2010; Kunnan, 2004; Spolsky, 2008). For this reason, the social implications of language testing were placed high on the agenda in the USA and United Kingdom (UK), raising issues such as ethics, the costs of high stakes tests, and the impact of failure on students' language learning identities (Davies, 2004; Shohamy, 2001; 2006; Spolsky, 2008). In this way, these language testing issues became critical debates that foregrounded both positive effects and concerns about the relevance of language tests and how it provides upward social mobility (Douglas et al., 2015).

Shohamy (1992, 2001) took up these critical issues and argued that centralised tests drive educational systems, and that language tests have power because they function as de facto language policies, especially in contexts where homogenous language policies exist. Subsequently, in the absence of democratic principles and practices, language tests are robust in "manipulating the linguistic repertoire, teaching material, language standards, and language diversity" (Shohamy, 2007; p. 126). Consequently, calls for democratic language tests paved the way for the emergence of critical pedagogy and critical language testing (CLT) to "limit, minimise, and control such powerful uses of tests" (Shohamy, 2001, p. 131).

CLT advocates call for increased investigations of tests' uses and their influences on education and society (Davies, 2004; Shohamy, 2001; Spolsky, 2008). In the UK, Australia, and the USA, research has found that tests narrow national curricula and teaching because teachers under pressure are prone to teach towards test items (Polesel et al., 2013; Stobart, 2008; Wyatt-

Smith et al., 2010). Furthermore, neoliberal testing practices that primarily originated in the USA and UK have become a global normative discourse that could unquestionably be adopted during educational reforms in developing countries, making schools more accountable for student performance (Klein & van Ackeren, 2011; Menken, 2008). For this reason, policymakers should know about the social and educational influences of testing practices because it can open up a space for valuable negotiations about the relationship between language policies and ways to mediate a more open, democratic, and socially-just culture of testing practices (Inbar-Lourie et al., 2013; Malone, 2013; Shohamy, 2021).

### **The multilingual turn and testing debates**

With the emergence of the multilingual turn, researchers raised another critical issue about assessing the linguistic competence of multilingual speakers (Canagarajah, 2006; Lopez et al., 2017; Shohamy, 2021). Before the multilingual turn, multilingual competence was based on the assumption that multilingual learners should achieve “native-like monolingual competence in each language” (Stavans & Hoffmann 2015, p. 157). For this reason, there is a growing interest in multilingual testing and assessment due to the transglobal and national flows of immigrant populations into mainstream education (Dahm & De Angelis, 2018; Cenoz & Gorter, 2020; Meier, 2017).

A few multilingual assessment research projects have emerged in various contexts, especially in countries experiencing high levels of multilingual diversity (Dendrinos, 2018). For example, the Multilingual Assessment in Education Project (MulAE) in Belgium experimented with the implications of computer-based science assessment in two languages (Turkish/Dutch or Polish/Dutch) compared to the evaluation in Dutch only (De Backer et al., 2019). In Israel, Shohamy (2011) advocates for similar multilingual testing (Hebrew only) or in two languages

(Hebrew and Arabic or Hebrew and Russian) for immigrant students whose home language is different from the official school language. In South Africa, Heugh et al. (2017) draw attention to how bi-/multilingual large-scale tests evaluate student knowledge in two or three languages. Multilingual assessment has also been advocated in New Zealand to address the needs of culturally and linguistically diverse immigrants (Sachtleben & Crezee, 2015). Therefore, several studies call for multilingual assessments so that multilingual students can engage in fair and equitable forms of testing that can contribute to “higher scores on academic tasks [that] more accurately reflects the knowledge of multilingual test takers’ (Shohamy & Menken, 2015, p. 421).

### **The emergence of language assessment literacy (LAL)**

The concept of Assessment Literacy (AL) encapsulates how different stakeholders' who make "important instructional and policy decisions" engage with assessment data at different levels in the educational system (Stiggins, 1991p. 535). In the field of language education, applied linguists have shown an interest in AL, but they argued for a language component to be added; thus, Language Assessment Literacy (LAL) emerged from Stiggins's (1991) original AL concept. In this way, LAL consists of language knowledge, language use in various contexts, and language teaching, which distinguishes it from general assessment literacy (AL) used in education (Bookhart, 2011; Popham, 2009). Therefore, LAL makes EFL teachers specialists because it is their unique form of professional understanding of the relationship between linguistic theories, the nature of language learning and teaching, and their engagement with language principles necessary for effective assessment (Giraldo, 2018).

Language assessment literacy (LAL) encapsulates three components which are the “knowledge, skills, and principles" associated with language learning and teaching (Giraldo,

2018, p. 180). Even though these three components have been at the centre of LAL debates and research, “its scope and boundaries have been questioned” (Giraldo 2018, p.180). For instance, two issues have been contentious and not explicitly resolved: 1) what language assessment literacy is and who should be considered as the stakeholders involved in LAL (Giraldo, 2018). Although there is an increased focus in the field on what LAL entails, there is agreement that the development of EFL teachers' LAL “is indeed necessary” (Giraldo, 2018, p. 180).

There have been robust debates about the stakeholders and participants that should have LAL knowledge in the LAL field. It is argued that besides language teachers, other stakeholders such as “school principals, parents, and politicians should know about language assessment and its implications” (Giraldo, 2018, p. 180). Moreover, Taylor (2013) argues that there is a need to shed light on the LAL of university administrators who evaluate high-stakes entry examinations such as IELTS, professional language testers, and test writers. In the South American context, LAL has been included in the professional development of experienced language teachers and pre-service language teaching degree programs (López & Bernal, 2009; Herrera & Macías, 2015). These initiatives in Columbia have shown the advantages for participants' sensemaking about the vital link between their language pedagogy and testing practices (Giraldo, 2018).

### **Empirical studies about AL and LAL**

The Assessment literacy of educators has been a prominent element of teacher education programs with growing attention to both large-scale (norm-referenced) and classroom assessments (DeLuca & Klinger, 2010). Unfortunately, many of the scholars who worked on measuring the assessment knowledge of teachers concluded that their AL was regrettably low (Alkharusi et al., 2011; Brookhart, 2011; DeLuca & Lam, 2014; Inbar-Lourie, 2008; Lam, 2015; Mertler, 2009; Popham, 2009; Scarino, 2013; Stiggins, 1991; Taylor, 2009; Vogt & Tsagari,

2014; Xu & Brown, 2017). In the USA, early research studies that measured teachers' assessment literacy through large-scale survey instruments highlighted the participants' limited knowledge and skills to assess student learning (Plake et al., 1993).

Alternatively, Campbell et al. (2002) administered an assessment literacy test to compare the knowledge of pre-service and in-service teachers and concluded that pre-service participants averaged two fewer correct answers than in-service teachers. However, later Mertler (2009) claimed that the similarity of scores among in-service and pre-service teachers supports the statement that most teachers do not have a sufficient level of assessment literacy. As a result, these studies pointed out the need to investigate teachers' and principals' perceptions and beliefs about assessment practices. Interestingly, the field of LAL shows even more complexity because EFL teachers need to include a much deeper engagement with 1) applied linguistics, 2) the nature of language learning and teaching, 3) second language acquisition theories (Davies, 2008; Giraldo, 2018, Scarino, 2013) and more recent components of multilingual assessment (Cenoz & Gorter, 2020; May, 2014, Meier, 2017).

In foreign language (FL) contexts, especially in contexts that face radical education reforms, language teachers' pedagogical and English language skills are receiving increased attention (Ilic & Warrington, 2007; Razmjoo, 2007; Tahereen, 2014). While these issues are essential, what is needed in these contexts is how teachers' LAL impacts their pedagogy and assessment because teachers' LAL can narrow the scope of learners' language learning (Chan & Evans, 2014). On the point of LAL in classroom assessment practices, López and Bernal (2009) found that Columbian teachers demonstrated divergent language assessment practices. For instance, their results showed that those teachers, who received language assessment training, were more able to develop assessment tasks that improved teaching and learning outcomes. In

contrast, teachers with no training only utilised summative assessment tasks to obtain grades. Therefore, this study concluded that teachers' beliefs and language assessment knowledge influenced their classroom assessment practices. As a result, teachers trained in LAL would probably include and use classroom assessment to draw conclusions about the academic achievements of their EFL students (Alkharusi et al., 2011; Brookhart, 2011; Popham, 2009; Shohamy, 2001; Stiggins, 1995).

In South America, research studies on EFL in-service teachers and pre-service student teachers AL concluded that these stakeholders need explicit training about LAL (Giraldo & Murcia, 2019). The study of Giraldo and Murcia (2018) illustrated how a LAL course for pre-service language teachers contributed to their development in two main areas. Their results indicated that the participants' language assessment views shifted from language assessment as tests that grade knowledge only to a process-oriented approach involving more critical discussions about language testing. In addition, Restrepo-Bolívar (2020) used learning journals to illustrate pre-service teachers' assessment perceptions and revealed divergent views of language assessment. Another study concluded that professional development programs for in-service teachers had increased their awareness of language assessment through reflection on topics such as the nature of language learning and second language acquisition pedagogies on language assessment (Giraldo & Murcia, 2019).

In the Asian context, it is noteworthy that LAL has been more concentrated in higher education contexts. For instance, Xu and Brown (2017) examined Chinese EFL university teachers' language assessment literacy using a modified variant of the Teacher Assessment Literacy Questionnaire (TALQ), produced by Plake and Impara (1992). The results showed that these participants' preliminary language assessment literacy levels were similar to findings in the

USA, Columbia, and Oman, recommending that universities provide regular professional development training on language assessment literacy.

Lam (2015) explored how five Hong Kong teacher education universities include (or not) language assessment literacy and development in pre-service teaching programs. The data included official websites of English Language Teaching Programs of the universities, government sources, and interviews with undergraduate students and instructors. The findings indicated a big theory-practice gap because the program content was too academic and did not primarily focus on developing critical LAL skills. For example, the study found that most teaching practices “primarily underscored the evaluation of pedagogical skills and classroom language proficiency rather than LAL” (Lam, 2015, p. 185). The researcher concluded that practicum supervisors assessed how future teachers delivered their lessons rather than how they evaluated their students’ language abilities. As a result, the study recommended that course-based language assessment training of pre-service teachers be re-designed to include LAL explicitly.

### **Summary**

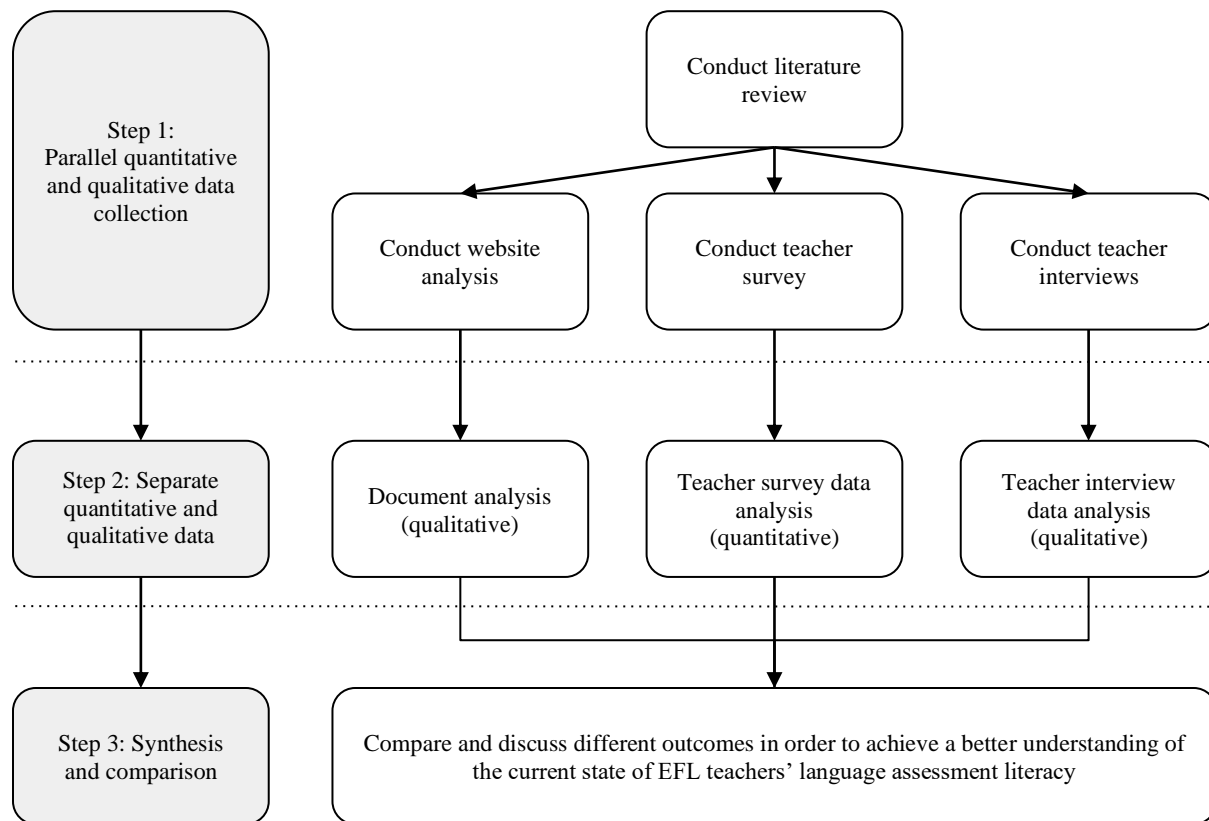
The chapter provided background about language testing in Kazakhstan and explained the development of LAL in various contexts such as the USA, South America, Western Europe, and Asia. The following section will present the study’s methodology.

### **Chapter 3: Methodology**

The research investigates how EFL teachers' LAL reflects the underpinnings of language assessment reform in the Kazakh context. For this reason, the purpose of this study was to shed light on how state-funded assessment PD courses impact EFL teachers' language assessment literacy. The previous chapter provided a review of LAL literature to illustrate its importance for EFL teachers. This chapter presents the research methodology and design and concludes with the study's ethical considerations.

#### **Research design and method**

Numerous studies have advocated for the dual benefits of collecting quantitative and qualitative data in educational research, referred to as mixed methods research (Creswell, 2014). One of the essential advantages of a mixed-method research design is that such studies build on the strengths of qualitative and quantitative approaches (Bryman, 2016; Creswell, 2014). For this reason, this study was underpinned by a mixed-method research design because several recently published educational research studies on assessment and assessment literacy advocate the advantages of following such an approach (Borg, 2015; Lam, 2015). For this reason, this study used a convergent parallel mixed-methods design. One of the strengths of a convergent parallel mixed-methods design is the possibility of combining complementary results from quantitative and qualitative research methods (Bryman, 2016; Creswell, 2014). As a result, it allowed the researcher to collect and analyse two separate datasets, the qualitative (website analysis and interviews) and the quantitative (online survey), and then merge them to compare and combine the results. Consequently, I used a parallel-databases variant where two parallel strands of data are gathered and examined independently and only brought together during the interpretation (See Figure 1) (Creswell, 2014).

**Figure 1***Overview of the Convergent Parallel Mixed-Methods Design*

*Note.* The figure shows the overall research grouped into three steps: eight main activities, where the actions after the literature review are conducted parallel and independently.

**Sampling and Research Instruments**

The sample consisted of one hundred and nineteen participants ( $n=119$ ) in a teacher survey, three participants ( $n=3$ ) in the one-to-one teacher interviews, and three English as a foreign language (EFL) teacher training course descriptions ( $n=3$ ) for document analysis. The study first used a convenience sampling strategy in which EFL teachers were invited to participate in an online survey through their corporate emails and the Kazakhstan Teacher of English Association webpage (<https://kaz-tea.kz/en/>). Finally, the study utilised purposive

sampling to select participants for the semi-structured interviews according to their teaching experiences and school locations.

***This study used the following research instruments:***

1. Documents (PD course descriptions from state-funded PD websites);
2. LAL (5-point Likert-type) Survey;
3. Semi-structured interviews.

The first research instrument consisted of documents and data on publicly accessible websites of state-funded PD organisations in Kazakhstan. It aimed to identify the content of EFL teachers' PD courses.

The second research instrument was a web-based survey for EFL teachers (see complete survey in Appendix A). The survey consisted of two parts: (1) a multiple-choice section to gather information about participants' teaching experience and participation in state-funded and other assessment PD interventions, and (2) a five-point Likert-type scale section which consisted of seventeen items aimed to assess teachers' perceptions about their assessment knowledge and skills. The survey items were adapted from a list of descriptions for LAL core components and their dimensions proposed by Giraldo (2018). To test for validity and internal consistency of data, the researcher computed Cronbach's alpha for each Likert-type item. It is recommended to have a minimum coefficient between 0.65 and 0.8 (or higher in many cases); coefficients that are lower than 0.5 are usually unacceptable (Bonett & Wright, 2015). As a result, Cronbach's alpha for all Likert-type items of the survey was higher than 0.8, which suggested a strong reliability measurement.

The third research instrument was semi-structured teacher interviews that consisted of six questions grouped into three categories, namely (1) warm-up questions about teachers'

experience in education and PD courses, (2) the questions about the teachers' attitudes and beliefs about language assessment in classrooms, and (3) the questions on the teachers' suggestions about language assessment in classrooms (see Appendix B). All the questions were tested in pilot interviews and checked for appropriateness to the research questions and the interview length. Overall, the main goals of the interview were:

1. To gain a deeper understanding of the teachers' LAL beliefs, attitudes, and skills in their classroom practices.
2. To gain a deeper understanding of the teachers' LAL perceptions and their suggestions about developing or improving their LAL.
3. To identify additional topics or issues that were not considered in the literature review and were potentially relevant to the study's research questions.

### **Data Collection Procedures**

After receiving the ethical approval from Nazarbayev University, Graduate School of Education (GSE) research committee, I sent emails with a link to the online questionnaire (Google Forms) to the intended organisations and school administrators of urban and rural public schools. First, I sent a follow-up email to explain the research purpose to two state-funded PD organisations to request assessment course documents and resources. Unfortunately, they were not forthcoming, and I then decided to collect course information freely available for public use on their websites.

Secondly, I sent the teachers an email explaining the study and provided a link to an interactive website (<https://sway.office.com/0ls9m24gjF1DAWF9>). It contained all the necessary research project information, such as the purpose, the ethical procedure to protect teacher participants' identities with the links to the survey, and the researcher's contact details. As a final

step in the survey, respondents were informed about the purpose of the follow-up semi-structured interview. I requested their willingness to participate through a tick list at the end of the survey questions.

Finally, I conducted online Zoom interviews at a convenient time for participants to ensure a comfortable atmosphere, which helped gain meaningful and reliable data (Creswell, 2014). The interviews lasted no more than 20 minutes; it was recorded after each participant gave consent.

### **Data Analysis Procedure**

Due to the study utilising convergent mixed-methods, the document analysis data, the quantitative data from the online surveys, and the qualitative data from the teacher interviews were collected simultaneously (Creswell, 2014). All the data was then analysed and reviewed to explain or support each other. Table 1 illustrates the study timeline.

**Table 1**

#### *Study Timeline*

<b>Activity</b>	<b>Date</b>
Proposal Presentation	October 2021
GSE Ethics Committee Approval	November 2021
Website analysis	November 2021
Participant Invitations & Online Survey Data Collection	November & December 2021
Conducting Interviews	January 2022
Data Analysis	February - April 2022

The analytical procedure for this research included document analysis, online teacher survey quantitative analysis, and qualitative analysis of teachers' interviews. Firstly, the researcher conducted a comprehensive study of the course descriptions and their overall structure for the website analysis to determine the knowledge and principles that underpinned the EFL

assessment PD courses in Kazakhstan. (See Appendix D). The screen captures of relevant Web pages were collected using the screenshot function of a computer from the desktop view of the websites. Relevant pages indicated the information about the courses for EFL teachers provided by state-funded PD organisations (See the screen captures in Appendix D). A second deductive analysis used Giraldo's (2018) LAL dimensions as a frame to evaluate the course learning outcomes and content (Giraldo, 2018).

The second phase of the study used SPSS software to provide descriptive and inferential statistics about the respondents' attitudes concerning their language assessment knowledge and the correlations between the completion of assessment courses and their self-perceived language assessment abilities. The researcher examined the reliability and internal consistency of data computing Cronbach's alpha and ran descriptive statistics to analyse demographic information and teachers' responses for Likert-type items (percentages, mean, standard deviation). In addition, Pearson's correlation was used to determine whether there were relationships between participants' completion of state-funded PDs, other assessment PDs and their perceptions about their own language assessment abilities and skills.

Finally, the teacher interview data were transcribed and coded for patterns emerging from the literature. These patterns were further analysed to identify recurring themes that revealed LAL results.

### **Ethical Considerations**

The researcher completed all necessary CITI Modules. Before collecting any data, the researcher first submitted the proposal for ethical approval. Upon approval, the researcher initiated the analysis of websites of PD course organisations and sent invitations to EFL teacher participants to participate in the online survey for a quantitative study. The invitation included

the study information, consent forms, a link to the online survey questionnaire, and the information on participation in individual semi-structured interviews.

The research included no more than minimal risk for the participants. Nonetheless, numerous precautions were used to protect the respondents who participated in this study.

As the proposed study utilised a mixed-method approach, it considered the ethical issues typical to the process. Ethics in mixed-methods studies need to relate to important quantitative and qualitative research issues. In sum, quantitative and qualitative research needs to be underpinned by ethics (Creswell, 2014).

### ***Confidentiality and Anonymity***

Researchers are responsible for ensuring that no harm is done to the participants at the initiation or during data collection (Bryman, 2016, Creswell & Clark, 2018). No harm means that an ethical researcher strives to ensure participants' anonymity and confidentiality. Therefore, throughout all stages of the study, I ensured that these ethical considerations were upheld by guaranteeing the confidentiality of all data and the anonymity of the respondents.

Although the data from the document analysis was taken from publicly available websites, the researcher removed identifiers of PD organisations, and changed names into codes such as Organisation 1 and Organisation 2. To ensure the privacy and anonymity of the survey participants, the online questionnaires were set up to make sure that participants' responses included minimal personal information and mainly focused on their educational background, knowledge of language assessment, and classroom practices. In addition, the account that was used in Google Forms had a strong-level password to guarantee confidentiality. Regarding the semi-structured interviews, the names of interview participants were changed to codes. The data they shared was coded, avoiding their personal information, and the recording started only after

their permission. Finally, all the research data was stored in Google Drive (a cloud-based storage solution) with a strong-level password that was only accessible by the researcher to ensure confidentiality and anonymity.

### ***Risks and Benefits***

Qualitative studies always involve risk; however, all the procedures were planned to mitigate potential risks. It was expected that teachers might feel uncomfortable talking about their thoughts on their professionalism, practices, or the quality of the professional development courses or programs they attended. To minimise these risks, the researcher explained the whole procedure of the study. Also, the participants were informed about the coding steps, data analysis, interpretation, and how the data will be used. I have also minimised the participants' fear of being recorded by explaining that the recordings would be deleted once the study was completed.

### **Summary**

Chapter 3 presented the research methodology, data collection methods, and analysis processes. The next chapter will present the study results that emerged from the convergent mixed-methods study.

## **Chapter 4: Findings**

The previous section focused on the research methodology to investigate EFL teachers' language assessment literacy related to the content of state-approved PD courses for in-service teachers. In this chapter, I present the results of a convergent parallel mixed-method study which included website analysis, a LAL survey, and teachers' semi-structured interviews. The central research question of this study was:

**To what extent do teachers' language assessment literacy (LAL) reflect the underpinnings of language teaching reform?**

To answer the main research question, I first need to present the findings from three sub-questions which were:

1. What is the content of language assessment PD training in Kazakhstan?
2. What are teachers' skills and knowledge about language assessment?
3. What are teachers' beliefs and attitudes about language assessment?

The study results are organised in the following sequence. First is document analysis, followed by the survey results, and the final section presents the interview findings.

### **Document Analysis**

The information from Organisation 1 and Organisation 2 about the courses for EFL teachers was analysed to illustrate the nature, focus, and content of EFL teachers' PD courses (See Appendix D for the original versions of the Russian content). The document sources used for this study are freely available on the publicly accessible websites of the state-funded organisations offering EFL PD courses. I now present the results: (1) Website Analysis and (2) Content Analysis.

### *Findings of the Website Analysis*

The first analysis phases searched for information related to the course goals, description, and overall structure to illustrate the underpinning assessment principles.

**Organisation 1.** The results show that the website of the Organisation 1 offers two courses for EFL teachers, titled *The advanced training of teachers at a secondary education school in the discipline "English" as part of updating the content of secondary education* (Course-1) and *Development and examination of assignments for assessment" in the subject "English language* (Course-2). Table 2 presents a grouped summary of Organisation 1's course overviews, their titles, delivery mode, overall course descriptions, and program goals.

**Table 2**

#### *Courses Provided by the Organisation 1*

<b>Course titles</b>	<b>The advanced training of teachers at a secondary education school in the discipline "English" as part of updating the content of secondary education (Course-1)</b>	<b>"Development and examination of assignments for assessment" in the subject "English language" Course-2)</b>
<b>Course mode</b>	Online, 3 weeks of training, including 160 academic hours of synchronous and asynchronous learning, 15 webinars of 90 minutes each day, 1 lesson plan for defence, and 1 certification presentation	Not mentioned
<b>Overall description and Goals</b>	To improve the professional skills of teachers in the context of updating the content of secondary education.	To improve the professional level of teachers in the field of development and examination assignments for assessment as part of updating the content of secondary education.
<b>Outcomes</b>	During the course teachers will: <ol style="list-style-type: none"> <li>1. understand the structure, content, goals and objectives of the updated English curriculum;</li> <li>2. know the features of the criteria-based assessment system, the practical aspects of its implementation;</li> <li>3. be ready to accept the methodology and content of the updated SES of general secondary education; to change the software and methodological support of the educational process, to change the goals and methods of pedagogical activity;</li> <li>4. learn how to build the educational process in accordance with the goals and expected learning outcomes of the Study Program.</li> </ol>	During the course teachers will: <ol style="list-style-type: none"> <li>1. study the main approaches to the development and examination of assessment tools;</li> <li>2. master the types and characteristics of tasks used in the assessment process;</li> <li>3. learn how to develop and conduct an examination of assessment tools;</li> </ol> Additional: In the process of coursework, special attention is paid to the development of different types of tasks, the ability to use different forms of their presentation, to improve pedagogical methods and approaches for examining the developed tasks, compiling test questions and distractors, the ability to analyse the content of tasks and test questions and select tasks for further use in the classroom in English.

**Organisation 2.** The website of the Organisation 2 has information on one course for EFL teachers, titled *Development of professional competencies of an English teacher (Course-3)* (See table 3). Table 3 summarises Organisation 2's course overviews, their titles, delivery mode, overall course descriptions, and program goals.

**Table 3**

*Course Provided by the Organisation 2*

<b>Course title</b>	<b>Development of professional competencies of an English teacher (Course-3)</b>
<b>Course mode</b>	Offline; or remote in asynchronous or synchronous mode, including 40 or 80 academic hours of learning in Russian, made for primary, basic secondary, general secondary education levels of education with a certificate as a completion form
<b>Overall description and Goals</b>	The purpose of the Program is to increase the level of professional competencies of English language teachers in the field in applying the achievements of pedagogical technologies, Internet learning for the development of "4C" skills of students in English lessons within the framework of the current Standard Curriculum for the subject "English Language."
<b>Outcomes</b>	After the course, teachers will: <ol style="list-style-type: none"> <li>1. carry out and optimise professional activities in accordance with regulatory legal acts in the field of implementation and development of teaching the subject "English";</li> <li>2. apply modern strategies for the formation of students' emotional intelligence to develop communication skills and speech activity of students;</li> <li>3. carry out effective interaction with students, based on the theory of 4Cs - the generation of such soft skills of students as communication, cooperation, critical thinking, creative skills;</li> <li>4. create and implement interactive tasks in the educational process with a focus on optimising the language skills of students;</li> <li>5. create differentiated tasks to involve all students in the educational process, including those with special educational needs;</li> </ol>

The analysis showed that both organisations focus on improving teachers' professional competencies; Organisation 1 noted that the program focuses on the content update for secondary education. In contrast, Organisation 2 pointed out that the program addresses the current Standard Curriculum for English. Course-1 (Organisation 1) and Course-3 (Organisation 2) are similar in their primary goals; both courses aim to improve EFL teachers' professional skills associated with the updated curriculum. The website revealed a few differences in the mode of delivery and the number of hours; 160 hours for Course-1 and 40 or 80 hours for Course-3 in offline and online modes. Also, Course-3 indicated its audience as the teachers of all levels of school education from primary to general secondary schools, whereas Course-1 and Course-2 did not specify their audience.

What is interesting in this data is that Course-2 is different from the other two courses because it focuses on developing teachers' assessment skills. Interestingly, the course description states that 1) special attention is paid to the development of different types of tasks, 2) the ability to use different forms of oral presentations, 3) to improve pedagogical methods and approaches for examining tasks, 4) compiling test questions and distractors, 5) the ability to analyse the content of tasks and test questions and finally 6) to select tasks for use in the English classrooms.

Overall, the results showed that Course-2 and Course-3 foregrounded the general teachers' competencies; thus, the focus on teachers' assessment literacy was vague. The analysis revealed that Course-2 has only one outcome about the features of criteria-based assessment systems and their implementation (See Table 2). Course-3 had two vaguely related assessment objectives: 1) creating and using interactive tasks, focusing on optimising students' language skills, and 2) creating and implementing differentiated tasks.

Finally, the outcomes of Course-2 were all related to assessment, which shows its focus was on improving EFL teachers' assessment knowledge. As shown in Table 2, Course-2 focuses on teaching the main approaches for developing and examining assessment tools, types and characteristics of tasks used in the assessment process and developing and examining assessment tools.

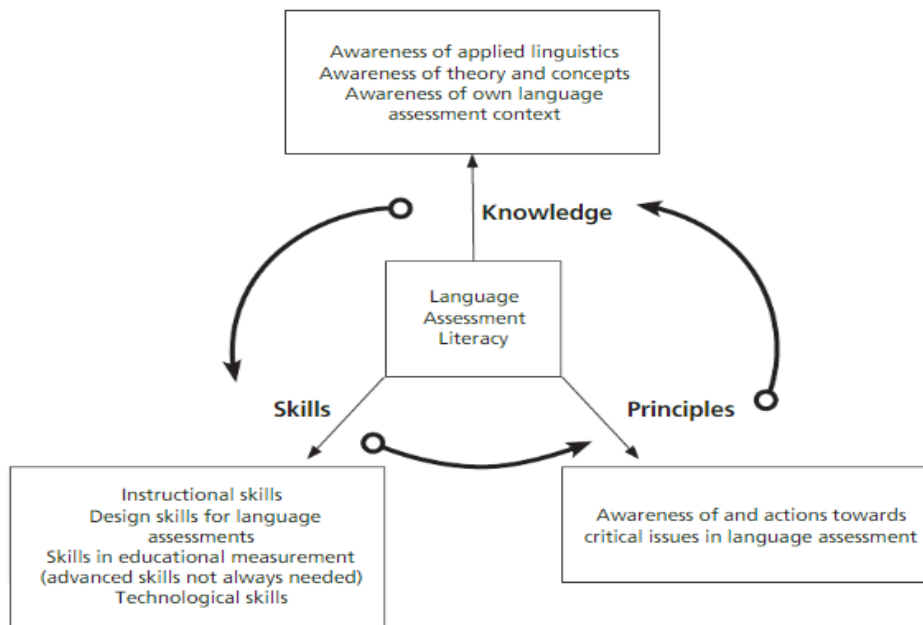
### ***Findings from Course Content***

The relation of courses' content to the features of LAL was analysed by categorising the course descriptions and outcomes against the language and assessment literacy (LAL) core list (Giraldo, 2018). This core list includes three central components: *Knowledge*, *Skills*, and

*Principles*, each with corresponding dimensions updated and developed with the knowledge of other researchers in the field (Davies, 2008) (See Figure 2).

## Figure 2

*A Core List of Language Assessment Literacy Dimensions: Knowledge, Skills, and Principles*



*Note.* Retrieved from Giraldo (2018, p. 187)

Comparing the website content on PD courses and the Core List, information on courses from the websites is not informative. Although Course-1 mentions studying the main approaches for developing and examining the assessment tools, no specific concepts, principles, or theories were mentioned. Similarly, all other information on Course-1, Course-2, and Course-3 did not include any special language assessment approaches. Therefore, it was challenging to discern whether the assessment component of PD courses makes connections with knowledge of applied linguistics issues such as communicative approaches to language assessment, second language acquisition and assessment, and assessment principles such as validity and reliability. For example, the description and outcomes of the Course-2 (see Table2) suggested that language

assessment is about “methods and approaches for examining the developed tasks” and “the ability to analyse the content of tasks and test questions and select tasks for further use in the classroom in English.” In addition, the course information suggested a limited focus on the relationship between instructional skills and the assessment of specific language competencies such as reading, writing, grammar, and listening and speaking. Instead, their focus was on teachers' knowledge about “the features of the criteria-based assessment system, the practical aspects of its implementation” (See Table 2, Course-1, outcome 2). Therefore, EFL teachers might gain a narrow perspective of language assessment in their professional development courses.

### *Summary of findings*

The results suggested that the content of both courses for EFL teachers only included assessment knowledge connected with the latest updated curriculum. For instance, the introduction of criteria-based assessment, summative and formative assessment methods, and approaches for selecting, developing, and examining different assessment tools, which in turn indicated that the PDs are working on improving teachers' teaching and assessment skills. However, the analysis also revealed that the courses lacked descriptions associated with applied linguistics such as second language acquisition, critical language testing (CLT), multilingualism, and assessment principles such as validity and reliability. These are all related to LAL's dimensions of knowledge and skills components (Giraldo, 2018). It suggested that teachers can probably conduct assessment procedures in a classroom setting; however, their knowledge might be insufficient to create and use high-quality assessment tasks and tools.

### **Quantitative Analysis**

This part of the thesis included a survey that consisted of two sections. The first section obtained demographic information about respondents' educational backgrounds, such as years of teaching, teaching categories, degrees, and whether they completed state-funded professional development courses, specifically language assessment. It was also used as a Qualifying demographic questionnaire to determine participant eligibility. The second section consisted of 17 Likert-type questions about EFL teachers' attitudes concerning their assessment competencies after educational reform. The questions were mandatory to avoid missing data, and it was based on the LAL descriptors proposed by Giraldo (2018). (See Appendix A).

### ***Presentation of Survey Findings***

This section presents the survey findings in the following sequence: (1) demographic information, (2) descriptive analysis of Likert-type items, (3) Pearson's correlation, and (4) the summary of findings.

### ***Demographics***

The first part of the survey revealed that a total of one hundred and nineteen teacher participants (n=119) completed the online survey, which consisted of 15 male (12.6%) and 104 female teachers (87.4%). The mean age of respondents was 34.12, and the mean number of years teaching was 10.50. The participants were also able to select their degree levels as one of the demographic questions on the survey instrument. The response options were BA (Bachelor of Arts), MA (Master of Arts), and Ph. D. (Doctor of Philosophy) or near completion. As a result, there was no dominant level of education, as the numbers of participants with Bachelor's and Master's degrees were almost similar, 56 (47.1%) and 55 (46.2%), respectively. The grade level taught was evenly distributed, with slightly fewer primary school EFL teachers than other grade

levels (16.9%). One-third of participants (70.6%) took state-funded PD courses on the updated curriculum, and more than half of them (52.1%) reported that they had other PD interventions on assessment. All demographic data collected from participants are presented in Tables 4 and 5.

**Table 4**

*Demographic Characteristics of Participants (N=119)*

Characteristics	<i>n</i>	%
Gender		
Male	15	12.6
Female	104	87.4
Location		
Rural	41	34.5
Urban	78	65.5
Highest level of education		
Bachelor's Degree	56	47.1
Master's Degree	55	46.2
PhD	8	6.7
Grade level taught		
Primary school (K-4)		
Secondary school (5-9)		
High school (10-11, and 12)		
Teachers who took state-funded PD training		
Yes	84	70.6
No	35	29.4
Teachers who had PD interventions on assessment		
Yes	62	52.1
No	57	47.9

**Table 5***Other Participant Characteristics (N=119)*

<b>Characteristics</b>	<b><i>M</i></b>	<b><i>SD</i></b>
Age	34.12	7.55
Years of teaching	10.50	7.25

Further, the participant responses were analysed using descriptive and inferential statistics.

### ***Quantitative Analysis of Likert-type items***

Responses for Likert-type scale items were quantified using a scale of one to five. One represented strongly disagree; two represented disagree, three represented not sure, four represented agree, and five represented strongly agree. To ensure that each participant's response was taken for each survey item, the section in the survey was mandatory.

Table 6 illustrates the correlation of respondents' age to their years of teaching. The years of teaching included the following ranges 1-3 years, 4-6 years, 7-12 years, 13-20 years, 21-31 years, and more than 32 years of experience. The participants' experience was used to determine the number of teachers who worked without any experience before the curriculum update, those who worked during the update, and teachers who had rich experience before the curriculum update. The descriptive analysis revealed that 16.0% of teachers started teaching when the reform was already launched, 22.7% of respondents worked during the planning stages of the updated curriculum, and 61.3% of respondents worked at schools with a traditional assessment model (see Table 6). It was anticipated that the teachers' perceptions about their language assessment knowledge and skills might differ according to the groups. However, the correlation analysis did not show any statistically significant correlation between grouped years of teaching

and responses to Likert-type item responses. Overall, the data showed that teachers' age and years of teaching were evenly distributed, and that teachers' years of experience did not affect their perceptions of their language assessment skills.

**Table 6**

*Years of Experience Correlated by Ages (N=119)*

Age	Year of experience					
	>3 years	4-6 years	7-12 years	13-20 years	21-31 years	<32 years
<b>21–25 years</b>	92.90%	7.10%				
<b>26–30 years</b>	15.60%	71.90%	12.50%	14.30%		
<b>31–35 years</b>	4.80%	14.30%	66.70%	53.80%	3.80%	
<b>36–40 years</b>			42.30%	61.50%	34.60%	3.80%
<b>&gt;40 years</b>						
<b>Total</b>	16.00%	22.70	24.40%	27.70%	8.40%	0.80%

Table 7 shows the teachers' self-reported attitudes toward their language assessment abilities, which were then correlated with the three core components of language assessment literacy: knowledge, skills, and principles, particularly the 17 descriptors selected for the Likert-type items (see Appendix A, Step 2). The survey results revealed a higher degree of positive responses for the *Principles* component, where 94.40% of respondents selected the “agree” and “strongly agree” values—the *Skills* component, especially instructional skills, with 93.08% of positive responses accordingly. The subgroup of items with the most negative responses was the *Knowledge* component, particularly the awareness of applied linguistics with 30.8% for the value “Not Sure” and 5.32% for “Disagree” and “Strongly Disagree,” respectively. Moreover, in

Figure 3, most of the negative responses for items from 10 to 14 indicated participants' responses to the *Knowledge* component in the LAL core list.

**Table 7***Attitudes of Teacher Survey Respondents to Their Abilities (N=119)*

Items	Responses					M* SD*	Cronbach's Alpha
	Strongly Disagree P*	Disagree P	Not Sure P	Agree P	Strongly Agree P		
<b>Instructional skills</b>							
1. plan, implement, monitor, record, and report student language development.	0.00%	0.80%	4.20%	57.10%	37.80%	4.32.596	.886
2. improve instruction based on assessment results and feedback.	0.00%	0.80%	1.70%	74.80%	22.70%	4.19.492	.885
3. utilize alternative means for assessment; for example, portfolios.	0.00%	0.00%	8.40%	67.20%	24.40%	4.16.552	.887
4. communicate norm- and criterion-referenced test results to a variety of audiences: students, parents, school directors etc.	0.00%	0.00%	11.80%	63.90%	24.40%	4.13.590	.889
<b>Design skills for language assessment</b>							
1. clearly identify and state the purpose for language assessment.	0.80%	0.00%	14.30%	63.90%	21.00%	4.04.656	.884
2. write selected-response items such as multiple-choice, true-false, and matching.	0.80%	0.00%	4.20%	70.60%	24.40%	4.18.577	.883
3. design rubrics for alternative assessments such as portfolios and peer-assessment.	0.80%	0.00%	18.50%	63.90%	16.80%	3.96.656	.881
<b>Technological skills</b>							
1. run operations on Excel; for example, descriptive statistics and reliability correlations.	0.00%	5.90%	26.10%	52.10%	16.00%	3.78.783	.891
2. use internet resources such as online tutorials and	0.80%	0.00%	4.20%	73.90%	21.00%	4.14.557	.885

adapt contents for their particular language assessment needs.

#### **Awareness of applied linguistics**

- |   |       |       |        |        |       |          |      |
|---|-------|-------|--------|--------|-------|----------|------|
| 1. compare approaches for language teaching and assessment; e.g., communicative language testing; task-based assessment.                                    | 0.80% | 5.00% | 28.60% | 57.10% | 8.40% | 3.67.738 | .881 |
| 2. explain the major issues in applied linguistics; e.g., bilingualism, language policy and planning, pragmatics, sociolinguistics, etc.                    | 0.80% | 3.40% | 37.00% | 50.40% | 8.40% | 3.62.725 | .879 |
| 3. analyse trends in second language acquisition and their impact on language assessment; e.g., motivation, cross-linguistic influence, learner strategies. | 0.80% | 5.00% | 26.90% | 61.30% | 5.90% | 3.66.704 | .882 |

#### **Awareness of theory and concepts**

- |  |       |       |        |        |       |          |      |
|--|-------|-------|--------|--------|-------|----------|------|
| 1. illustrate the history of language testing and assessment, and its impact on current practices and society. | 0.00% | 5.00% | 37.80% | 52.10% | 5.00% | 3.57.671 | .885 |
| 2. evaluate the kind of washback that assessments can have on learning, teaching, curricula, and institutions. | 0.00% | 6.70% | 36.10% | 47.90% | 9.20% | 3.6.751  | .890 |

#### **Awareness of and actions towards critical issues in language assessment**

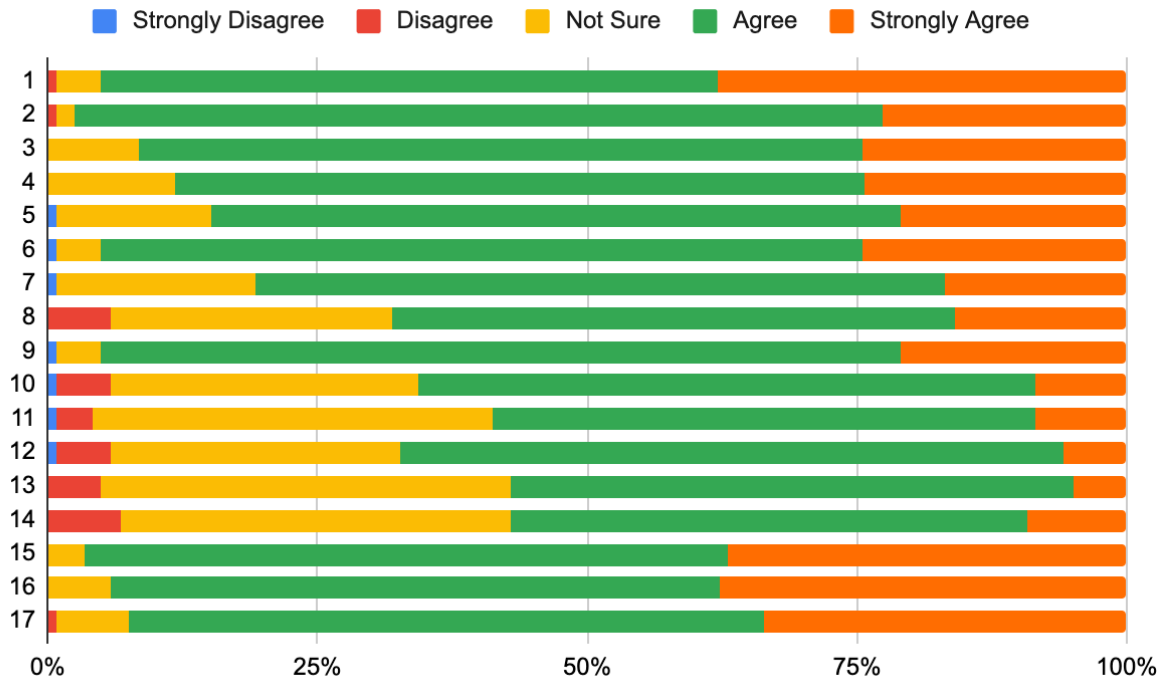
- |   |       |       |       |        |        |          |      |
|---|-------|-------|-------|--------|--------|----------|------|
| 1. use tests, test processes, and test scores ethically.  |       |       |       |        |        |          |      |
| 2. provide assessment practices that are fair and non-discriminatory.   | 0.00% | 0.00% | 3.40% | 59.70% | 37.00% | 4.34.541 | .888 |
|   | 0.00% | 0.00% | 5.90% | 56.30% | 37.80% | 4.32.581 | .890 |
| 3. implement democratic language assessment practices, by giving students opportunities to share their voices about assessment. | 0.00% | 0.80% | 6.70% | 58.80% | 33.60% | 4.25.614 | .885 |

---

P\* = percentage    M\* = mean    SD\* = standard deviation

**Figure 3**

Bar Chart For Attitudes of Teacher Respondents to Their Abilities (N=119)



*Note.* This is a stacked bar chart where the rows represent 17 items of Likert-type scale questions, where row one presents the results for the first item in the second part of the questionnaire titled: *The current state of EFL teachers' LAL* in Appendix A, step 2.

Moreover, as seen in Table 7, the participants most agreed with the items two and 15 (97.5% and 96.7%), which indicated that they can “improve instruction based on assessment results and feedback” and that they find it easy to “use tests, test processes, and test scores ethically.” The items participants most disagreed with were questions 14 and eight (6.7% and 5.9%). This indicated that they find it easy to “evaluate the kind of washback that assessments can have on learning, teaching, curricula, and institutions” and that they can “run operations on Excel; for example, descriptive statistics and reliability correlations.” In addition, the analysis found that items 13, 11, and 14 (37.8%, 37%, and 36.1%) had the most responses for the value

“not sure.” In sum, the participants reflected that they are more prepared to use assessment data to improve their teaching and to do it ethically.

Also, the results suggested that the respondents lacked the knowledge related to the significant issues in applied linguistics, the history of language testing and its impact on current practices and society, and the evaluation of the kind of washback that assessments can have on learning, teaching, curricula, and institutions (Giraldo, 2018).

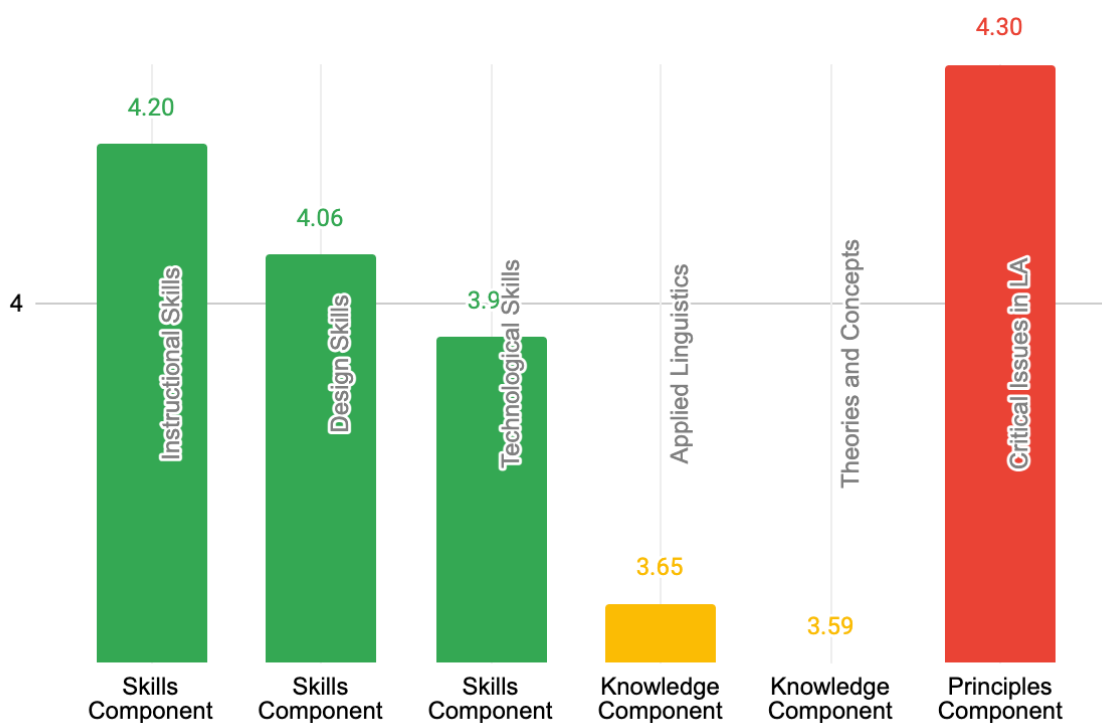
The results have four well-defined suggestions about teachers' knowledge, skills, and principles in language assessment. What was evident from teachers' responses was that they seemed to have more confidence in using language assessment and test processes ethically and providing fair and non-discriminatory assessment practices. Another significant finding was teachers' responses about their knowledge of Skills component items. The teachers were particularly confident in Instructional skills that included recording and reporting assessment data, using it for improving instructions, and utilising different types of assessment. Also, they reflected that they were mostly good at designing the assessment tools, identifying their purpose, and preparing the task rubrics. However, teachers had low responses for item eight (see Appendix A, step 2) about running operations on Excel for measurement tasks like descriptive and correlations analysis of assessment data. In addition, another significant finding was that teachers indicated limited awareness of applied linguistics concepts associated with language testing.

Figure 4 represents the log-transformed column chart of average means by LAL dimensions as a visual reference. These dimensions include instructional skills, design skills, and technological skills of the Skills component, the awareness of applied linguistics and the awareness of theory and concepts of the Knowledge component, and the awareness of and

actions towards critical issues in language assessment of the Principles component. As seen in Figure 4, the average means of the responses point to the low scores of the dimensions in the Knowledge component. Means 3.65 and 3.59 are close to the value “not sure” of the Likert-type responses.

**Figure 4**

*Log-Transformed Column Chart for Average Means by LAL Dimensions*



*Note:* The grid line with the number four on the horizontal axis represents the value “agree” on the Likert scale.

**Pearson’s Correlation.** The researcher evaluated the relationship between the content of PD courses and teachers' assessment abilities to answer the main research question: To what extent do teachers' language assessment literacy (LAL) reflect the underpinnings of language teaching reform?

The researcher correlated EFL teachers’ completion of state-funded courses or other language assessment PD interventions against their responses on language assessment literacy's

three components (knowledge, skills, and principles). The aggregate totals from participants' responses were used to perform calculations based on the components that respondents selected as being capable or not capable. The Skills component consisted of nine items, the Knowledge component had five items, and the Principles component had three items, requiring respondents to select a choice on a five-option in Likert-type scale (strongly disagree, disagree, not sure agree, or strongly agree).

To quantify the level of the teachers' abilities in Skills (LAL-S), Knowledge (LAL-K), and Principles (LAL-P) components (see Tables 8 and 9), each strongly disagree was assigned a value of one, disagree a value of two, not sure a value of three, agree for a value of four and strongly agree received a value of five. LAL-S, LAL-K, and LAL-P aggregate scores were generated as the dependent variable for each respondent who answered the questions. The independent binary variables were the Completion of state-funded courses (SFC) and Completion of other PD interventions (PDI), where *yes* was coded as 0 and *no* as 1.

A Pearson's correlation was run to test the relationship between the independent variables of LAL components aggregate scores (LAL-S) (N = 119, M ~ 39.9, SD ~ 3.744) (LAL-K) (N = 119, M ~ 18.1, SD ~ 2.998), (LAL-P) (N = 119, M ~ 12.9, SD ~ 1.495).

Table 8 summarises the descriptive statistics of these variables. As seen in Table 9, data analysis reveals a statistically significant weak negative relationship between the completion of PD interventions and their abilities in the Skills component ( $r = -.218$ ,  $n = 119$ ,  $p = .009$ ). A one-tailed test was chosen because of the underlying assumption in the research hypothesis that the direction of a teacher's ability scale scores (higher or lower) will depend on their SFC or PDI scores. The result was interpreted as the variables having a correlation of 0.218, which was significantly below the 0.01 level. Results also indicate the 0.218 correlation relationship due to

chance at a level of  $p < 0.01$ , which was even a stronger finding than the acceptable level of significance set for the study requiring a p-value where  $\alpha < 0.05$ .

**Table 8**

*Descriptive Statistics for Dependent Variables Scores and Independent Variables Aggregate Scores*

Variables	N	Minimum	Maximum	Mean	SD
Dependent					
SFC	119	0	1	.29	.458
PDI	119	0	1	.48	.502
Independent					
LAL-S	119	16	45	36.9	3.744
LAL-K	119	10	25	18.1	2.998
LAL-P	119	9	15	12.9	1.495

**Table 9**

*Correlation between Course Completion and LAL Scores*

		LAL-S	LAL-K	LAL-P
SFC	Pearson Correlation	-.027	.065	.065
	Sig (1-tailed)	.385	.240	.242
	N	119	119	119
PDI	Pearson Correlation	-.218**	-.148	-.110
	Sig (1-tailed)	.009	.055	.117
	N	119	119	119

\*\* . Correlation is significant at the 0.01 level (1-tailed)

Correlation analysis was further used in an attempt to determine if there were relationships between the PDI and LAL-S sub-groups, specifically the Instructional skills (LAL-S-I), Design skills (LAL-S-D), and Technological skills (LAL-S-T) components. The completion

of PDI as self-reported by teacher participants shows a statistically significant negative weak relationship with the Instructional skills ( $r = -.192$ ,  $n = 119$ ,  $p = .018$ ) and Design skills ( $r = -.213$ ,  $n = 119$ ,  $p = .010$ ) (Table 10, Table 11). Again, each correlation coefficient is negative, indicating that as the score for the PDI completions increases, the score for the Instructional skills and Design skills or the variables would decrease. This negative correlation supports the directional hypothesis associated with the research question, which means that the teacher respondents who indicated that they had other PD interventions on assessment were more confident in their knowledge and skills in LAL components.

**Table 10**

*Descriptive Statistics for PDI and LAL-S Sub-groups Aggregate Scores*

Variables	N	Minimum	Maximum	Mean	SD
PDI	119	0	1	.48	.502
LAL-S					
LAL-S-I	119	10	20	16.8	1.720
LAL-S-D	119	3	15	12.2	1.592
LAL-S-T	119	3	10	7.92	1.105

**Table 11**

*Correlation between PDI and LAL-S Sub-scores*

		LAL-S-I	LAL-S-D	LAL-S-T
PDI	Pearson Correlation	-.192**	-.213**	-.133
	Sig (1-tailed)	.018	.010	.075
	N	119	119	119

\*\* . Correlation is significant at the 0.05 level (1-tailed)

*Overall, the results revealed the following well-defined trends in descriptive analysis:*

1. The survey results revealed that the teacher respondents indicated themselves as being more confident in using language assessment and test processes ethically and providing fair and non-discriminatory assessment practices.
2. The teachers also had higher responses for recording and reporting assessment data, using it to improve instructions, and utilising different assessment types.
3. The respondents seemed less familiar with the functions of Excel or with terms such as descriptive statistics and reliability correlations.

*The following trends in the correlation analysis,*

1. The results showed that the teachers' LAL skills related to the completion of PD interventions teachers' LAL skills ( $r = -0.22$  and  $-0.15$ ,  $p < 0.05$ ). Thus, the teachers who individually access the assessment interventions other than state-funded ones seem to have more confidence in assessment knowledge and skills.
2. In addition, there was no statistically significant correlation between state-funded PD courses and teachers' LAL, which means that teachers' completion of the courses might not have much influence on their assessment abilities.
3. Within the Skills component, the completion of PD intervention is related to teachers' abilities in Instructional and Design skills ( $r = -0.19$  and  $-0.21$ ,  $p < 0.05$ ). Each of the correlations was negative, which means people who indicated no experience in PD interventions seemed to have lower abilities according to their self-reported responses.

As per the conventions associated with the convergent–parallel mixed-methods approach, the next section will present the results from the qualitative interviews.

### **Presentation of Qualitative Data**

The following section presents the findings from interviews with three participants.

I selected three respondents who indicated their willingness to participate in a follow-up interview from the questionnaires. The purpose of the interview was to give the researcher insight into the teachers' attitudes and beliefs about language assessment in their classroom practices. Also, I wanted to shed light on their opinions about how the assessment-related PD courses informed their assessment beliefs and practices.

### ***Presentation of the Interview Results***

The interview consisted of six questions to illustrate teachers' attitudes and beliefs about classroom-based language assessment. The questions were divided into three categories: 1) teachers' background in education, 2) their attitudes and beliefs about language assessment, and 3) their preferences about types of interventions for professional development and suggestions about the ways to improve the teachers' language assessment knowledge and skills (See Appendix B for the list of questions in the semi-structured interview).

Purposive sampling was used to select the teacher participants according to their teaching experiences from three different perspectives. The sample consisted of three EFL participants: Participant 1 teaches at a rural school but previously worked at an urban school. Participant 2 is an experienced teacher at an urban school. Participant 3 is a pre-service EFL teacher still at university, completing her EFL teaching degree.

### ***Participants' Profile***

Before the interview, the respondents were asked to complete a brief survey about their demographic and background information, including their teaching (years) experience and the types of PD interventions they participated in (see Table 12). The data revealed that the

participants' teaching experience ranged from 0 to 7 years and that their age ranges were between 23 and 35. In addition, two participants completed state-funded PD courses on curriculum reform.

**Table 12**

*Personal Background of Interview Participants (N=3)*

Questions	Participant 1	Participant 2	Participant 3
Gender	F	M	F
Age	26	35	23
Category	In-service	In-Service	Pre-service
Type of school	Rural	Urban	
Year of teaching	7	6	
State-funded PD course	Yes	Yes	No
Worked at school before the curriculum update reform	Yes	No	lower than medium
Self-perceived LAL level	middle	lower than medium	

### *Teachers' self-perceived assessment knowledge*

The participants were asked to share their thoughts about their assessment beliefs to understand the teachers' beliefs and attitudes about their assessment knowledge and practices. In addition, the respondents were asked to indicate their self-perceived language assessment knowledge and explain their answers (See Appendix B, Questions 4, 5).

The interview analysis revealed a dominant theme associated with teachers' average self-perceived knowledge about assessment. For example, Participant 1 said,

Regarding my own assessment, I can say that I am somewhere in the middle because we are teachers at an ordinary rural school.

Also, Participant 2, an urban schoolteacher, indicated her self-perceived LAL as average. She stated,

I think my assessment is a bit lower than medium.

Participant 3, a pre-service teacher, also said,

Maybe, it is about the normal or under the normal.

In addition, one participant raised English language proficiency and the urban-rural divide as factors that impacted teachers' assessment abilities. For example, Participant 1 believes that teachers' English language proficiency and geographical location as factors influencing classroom assessment practices,

It can be challenging for a teacher to assess a student's performance when the teachers themselves have poor English. Another factor is the lack of attention to teachers from rural schools. I have worked here for three years and did not participate in any seminar or workshop.

Participant 3 believed that assessment theories do not relate to classroom-based challenges. He commented,

All the knowledge, even the theories from books, are not always practical, there are various situations in classrooms, and it is hard to predict them and act as it is instructed in documents [and PDs].

Participant 2 believed there is a gap between knowing about assessment and applying it in practice. He said,

Even though I know a lot about criteria-based assessment and some other basics, mainly due to lack of time, the process of assessment happens in the form of a quick review.

To sum up, data showed that despite their differences in experience and locations, all of their self-perceived LAL perceptions were average; thus, their self-perceived LAL was similar to recent large-scale surveys that found teacher assessment literacy was regrettably low (DeLuca & Lam, 2014; Vogt & Tsagari, 2014; Mertler, 2009; Lam, 2015; Xu & Brown, 2017).

Unfortunately, it holds implications for teachers and learners because a teacher's average LAL can limit pedagogy and narrow learners' scope of language learning and developing communicative competence (Chan & Evans, 2011). Additionally, the participants' LAL beliefs showed English language proficiency, geographical locations of schools, and that knowledge of assessment does not help in practice. As a result, their self-perceived LAL and LAL beliefs implied that they would probably be more comfortable utilising summative assessment tasks to obtain grades despite PD training (López & Bernal, 2009).

### *Managing assessment in practice*

The interview revealed that the participants' self-perceived LAL and LAL beliefs could affect their classroom assessment practices. A dominant theme was managing assessment in their schools or classes. For example, Participant 1 said,

Imagine that only one teacher is responsible for teaching all primary classes, more than 100 students overall. So, the quality of teaching [and assessment] suffers.

Participant 3 noted,

Teachers should inform students about skills and knowledge to be acquired by the end of a course [but] teachers should be [knowledgeable about] the steps and actions for some unexpected situations [for example] teachers [should] know that there are different students with different needs.

Another factor emerging from managing assessment in EFL classrooms was relying on guidebooks when developing assessment tools. Participant 1 noted,

I take most of the exercises from a book specially designed for a specific grade, e.g., Excel for grade 5, Smiles for primary school, so I can be at least 60% sure that my assessment is fair.

It was also echoed in the responses of Participant 3,

If I am provided with teaching and assessment materials with tasks and instructions especially prepared for specific school age, I think I would be 80-90% sure that I teach and assess in the right way.

In summary, the extracts indicated teachers' workload and low confidence in developing assessments, indicating gaps in their pedagogical and technical knowledge of language assessment. All the participants voiced a lack of confidence in their ability to develop assessments without guides or textbook resources. As a result, despite attending assessment PDs, the participants in this study would probably possess limited capacity to draw conclusions about the academic achievements of their EFL students (Popham, 2009; Alkharusi et al., 2011). For instance, participants' responses did not demonstrate 1) "an increased awareness of and appreciation for assessment as a tool for guiding and improving language instruction" (Montee et al., 2013, p. 23) or 2) "a process in which the teacher gathers relevant information about the student's weaknesses and strengths in the learning process to make decisions about the instruction and students' learning" (Restrepo-Bolívar, (2020, p. 45).

### ***Professional Development and EFL Teachers' Professional Identity***

During the interview, the participants were asked to reflect on their assessment knowledge and its influence on their professional identities. Participant 3 noted,

It has a significant impact on a teacher's work at school because parents need to know the level of their children, the school needs to see the teacher's work, and the students want to know about their progress, so maybe this is the reason for schools to have such rating system to assess everything.

The above extract indicates that the teachers' professional identity was related to accountability associated with reporting summative assessment results to the school administration, parents, and students.

Interestingly, Participant 2 mentioned the importance of understanding summative and formative assessments. He responded,

If a teacher is professionally informed on the assessment system, it will generally help them check summatives and formatives or other works, explain their marks to students and other colleagues, why and how they have a specific rating, and

improve the teaching. So, teachers and students could work in tandem when the process is more evident to both, and its impact is only positive.

Another significant issue raised about participants' professional identity was making assessment information, presentations, and documents practically relevant and usable. On this issue, Participant 2 commented,

We annually have different contests at a national level for teachers. So, we have winners, and maybe they could produce valuable materials, not just academic articles, which can be boring for the majority, but maybe some exciting presentations, of course, unobtrusive, lite, with elements of humour, focusing on the masses. It could be beneficial.

Similarly, Participant 3 said,

We had different reforms and changes in the educational system, but they were not correctly presented, so teachers misinterpreted them, and finally, students faced problems.

The extracts illustrated that teachers' LAL professional identities after PD training mostly foregrounded summative assessment and accountability of reporting results to school administration and parents. It was illuminating that their professional identities did not make explicit their knowledge about theories and models of language proficiency, their educational measurement designs, and the impact of language testing on their classroom-based practices (Davies, 2008). Another dominant theme was that the assessment materials and initiatives to improve EFL teachers' LAL was theoretical and not accessible to improve classroom-based assessment practices. The participants' responses suggested they were probably not exposed to analysis or design of assessment instruments (Kleinsasser, 2005; Restrepo-Bolívar, 2020). As a result, it reflected other studies where teachers' expectations were about having practical, first-hand training (Fulcher, 2012; Giraldo & Murcia, 2018). This finding emphasised that if teachers study LAL knowledge and principles, they should do so within a practice-based framework.

*Developing teachers' language assessment skills for classroom-based practices*

Participants foregrounded the importance of working in collaboration and constantly sharing ideas among colleagues from different schools to improve their LAL. For instance, Participant 2, from an urban school, noted,

The most optimal variant [is] listening to your colleagues who have already mastered [assessment] to a certain level. And the views of teachers from other places are also essential. We should combine the knowledge from external sources and local practices.

The pre-service teacher (university student) indicated that having knowledgeable mentors could be helpful for novice teachers' LAL development. Participant 3 said,

Well-prepared mentors at schools who are informed and experienced in most schoolwork could be helpful. They could advise or guide new teachers, and it would help to learn fast and effectively.

Participant 2 commented on minimising teachers' workload and simplifying reporting on assessment outcomes. He said,

The important thing is to find an appropriate time and not to bring more work to teachers. For example, additional reports [which] should have a more accessible and concise format, brief, only containing the primary information. As a result, every teacher may have their own notes or a handbook, and maybe some practical exercises that are not time-consuming.

In conclusion, these extracts showed a common view about effective intervention for developing EFL teachers' LAL. The first theme was the value of teachers' collaboration across schools and geographical locations. Therefore, these responses indicated more importance in collaboration with more knowledgeable and experienced peers. Their responses indicated that scaffolding their LAL development would best occur by sharing classroom-based assessment experiences contingent upon relationships with mentors, fellow novice teachers, and interaction with experienced teachers (Vélez-Rendón, 2006). As a result, these extracts indicated a positive stance toward creating conditions for the co-construction of assessment knowledge and

understanding rather than PD sessions that add to teachers' workloads associated with cumbersome institutional assessment reporting.

### *Summary of interview findings*

First, the interview data showed that despite their differences in experience and locations, the participants' self-perceived LAL perceptions were average; thus, their self-perceived LAL was similar to recent large-scale surveys that found teacher assessment literacy was regrettably low (DeLuca & Lam, 2014; Vogt & Tsagari, 2014; Mertler, 2009; Lam, 2015; Xu & Brown, 2017). As a result, their self-perceived LAL and LAL beliefs implied that they would probably be more comfortable utilising summative assessment tasks to obtain grades despite PD training (López & Bernal, 2009).

Secondly, the data illustrated teachers' workload and low confidence in developing assessments, indicating pedagogical and technical knowledge gaps. Despite attending assessment PD courses, they were over-reliant on assessment guides, which highlighted the participants' limited capacity to draw conclusions about the academic achievements of their EFL students (Alkharusi et al., 2011; Popham, 2009).

The third finding illustrated that after PD training, the teachers developed LAL professional identities related to summative assessment and institutional accountability of reporting results to school administration and parents. For example, their LAL professional identities did not include core language skills such as being more aware of assessing listening, speaking, reading, and writing. As a result, the results demonstrated that teachers might not shift their assessment practices from grammar and vocabulary to assessing communicative competence optimally.

Finally, the study found that teachers indicated a positive stance toward creating conditions for the co-construction of assessment knowledge and understanding rather than PD sessions that add to teachers' workloads associated with cumbersome institutional assessment reporting. Therefore, collaboration across schools and geographical locations was viewed as the best way to develop teachers' LAL.

### **Summary of findings across the three data sets**

The analysis of the websites suggested that PD courses funded by the government tend to focus less on fostering the development of LAL and more on the general improvements of the professionalism of EFL teachers. Although they include some knowledge of language assessment, identifying their features and comparing them to the core components of LAL was challenged due to the lack of information on specific principles or concepts of language assessment in the course descriptions on the websites.

The quantitative study revealed three well-defined trends. First, the results showed that the teacher's completion of other PD interventions related to teachers' LAL skills ( $r = -0.22$  and  $-0.15$ ,  $p < 0.05$ ). Second, and in contrast, there was no statistically significant correlation between state-funded PD courses and teachers' LAL, meaning that completing the PD courses did not influence teachers' perceptions about their LAL abilities. Third, within the Skills component, the completion of PD intervention was related to teachers' Instructional and Design skills ( $r = -0.19$  and  $-0.21$ ,  $p < 0.05$ ). Each correlation was negative, meaning people who did not attend PD interventions seemed to have lower abilities according to their self-reported responses.

Finally, the qualitative study results indicated that 1) the participants' self-perceived LAL perceptions were average, 2) gaps in teachers' pedagogical and technical knowledge resulting in low confidence in developing assessments, 3) the teachers developed LAL professional identities

related to summative assessment and institutional accountability and 4) the participants held a positive stance toward collaboration about the co-construction of assessment knowledge rather than PD sessions. See Table 13 for a visual summary of the findings of this convergent mixed-method study.

**Table 13***Joint Display of Quantitative and Qualitative Summarised Findings*

LAL Core Components (& Dimensions)	Findings		
	Document Analysis (Qualitative) (n=3)	Teacher Survey (Quantitative) (n=119)	Teacher Interview (Qualitative) (n=3)
<b>Skills</b> (Instructional skills, design skills for language assessment, and technological skills)	No data	Agree*: 88.43% Mean** : 4.10 <ul style="list-style-type: none"> <li>• The second highest dimension, Instructional skills (Agree*=93.08%, M**=4.20)</li> <li>• The highest item was the ability to <i>plan, implement, monitor, record, and report student language development</i> (Agree=97.50%, M=4.32)</li> <li>• The lowest scored item within the component was in Technological skills dimension, the ability to <i>run operations on Excel; for example, descriptive statistics and reliability correlations</i> (Agree*=68.10%, M=3.78)</li> </ul>	The participants' self-perceived LAL perceptions were average.
<b>Knowledge</b> (awareness of applied linguistics, theory, and	No data	Agree*: 61.14% Mean** : 3.62 <ul style="list-style-type: none"> <li>• The component with most negative responses</li> </ul>	The participants' self-perceived LAL perceptions were average.

concepts)		<ul style="list-style-type: none"> <li>Specifically, the awareness of theory and concepts dimension (<math>M^{**}=3.58</math>) with 36.95% for the value “Not Sure” and 5.85% for “Disagree” and “Strongly Disagree.”</li> </ul>	
<b>Principles</b> (awareness of and actions towards critical issues in language assessment)	No data	<p>Agree*: 94.40% Mean**: 4.30</p> <ul style="list-style-type: none"> <li>The highest overall score for the component and the dimension</li> <li>The highest item, <i>using tests, test processes, and test scores ethically</i> (Agree=96.70%, <math>M=4.34</math>)</li> </ul>	The participants' self-perceived LAL perceptions were average.
	Additional: PD courses lack focus on EFL teachers' assessment knowledge skills	<p>Additional: EFL teachers' LAL skills related to the completion PD interventions teachers LAL skills (<math>r = -0.22</math> and <math>-0.15</math>, <math>p &lt; 0.05</math>). However, there was no statistically significant correlation between state-funded PD courses and teachers' LAL.</p> <p>Within the Skills component, the completion of PD intervention related to teachers' abilities in Instructional and Design skills (<math>r = -0.19</math> and <math>-0.21</math>, <math>p &lt; 0.05</math>). Each of the correlations was negative, which means people who did not attend PD interventions seemed to have lower abilities according to their self-reported responses.</p>	<p>Additional: First, the participants' self-perceived LAL perceptions were average.</p> <p>Secondly, teachers demonstrated low confidence in developing assessments, indicative of pedagogical and technical knowledge gaps, despite attending assessment PD courses.</p> <p>Third, teachers' beliefs and values foregrounded summative assessment and institutional accountability. It was illuminating that their professional identities did not make explicit the relationship between models of language pedagogy and assessment design, their educational measurement designs, and the impact of language testing on their classroom-based practices.</p> <p>Finally, the study found a positive</p>

			teachers' stance toward creating conditions for the co-construction of assessment knowledge collaboration and mentoring rather than attending more PD sessions.
--	--	--	---

Note: \*. Combined percentages for values “agree” and “strongly agree” (see more detailed in Table X); \*\*. Average score for mean

**Conclusion**

This mixed-methods research study aimed to investigate how EFL teachers' LAL reflects the underpinnings of language assessment reform in the Kazakh context. For this reason, the purpose of this study is to shed light on how state-funded assessment PD courses impact EFL teachers' language assessment literacy. This chapter presented the results from documents, quantitative survey analysis and semi-structured interview qualitative analysis with three teacher participants. The following chapter will present the discussion of the findings.

## Chapter 5: Discussion

The study investigated the LAL of EFL teachers in Kazakhstan to determine how their assessment knowledge and skills correspond to the current curriculum updates for language pedagogy and assessment. The previous chapter presented the findings from the document analysis, a teacher survey, and interviews. This chapter discusses the results in relation to the literature on assessment literacy, language assessment literacy, and teachers' perceptions about language assessment to answer the main research question of the study:

**To what extent do teachers' language assessment literacy (LAL) reflect the underpinnings of language teaching reform?**

To answer the main RQ, I will first address the sub-questions, which were:

1. What is the content of language assessment PD training in Kazakhstan?
2. What are teachers' skills and knowledge about language assessment?
3. What are teachers' beliefs and attitudes about language assessment?

**Research Question 1: What is the content of language assessment PD training in Kazakhstan?**

The study found that the state-funded PD organisations' course objectives are working towards improving EFL teachers' professional skills, including assessment knowledge, which was visible in the information on their websites. However, the focus was on developing general pedagogical knowledge and introduced criteria-based assessment with primary emphasis on summative assessment and formative assessment strategies. First, the results indicated that significant LAL components were not covered in these courses. Therefore, these courses mostly omitted significant LAL dimensions necessary for effective language testing. These dimensions were (a) essential knowledge of applied linguistics (e.g., communicative approaches of language

teaching) and their contexts for language assessment (e.g., ability to be critical towards assessment practices they use); (b) instructional and technical skills about the relationship between pedagogy and assessment design, creating assessment tools for the four language skills, and basic knowledge of educational measurements; and (c) the social consequences of assessment concerning issues of power and ethics and fairness in assessment practices (Davies, 2008; Fulcher, 2012; Giraldo, 2018; Inbar-Lourie et al., 2013). Therefore, these gaps in the PD courses suggested that EFL teachers in Kazakhstan might lack confidence in creating high-quality assessment data.

The second finding showed a lack of course content about the knowledge component of LAL. This was also confirmed in the quantitative and qualitative data analyses. For example, in the teacher survey, the dimensions in the knowledge component had the lowest mean of teachers' responses (average mean for Applied Linguistics = 3.65, and Theories and Concepts = 3.59). The interview showed the respondents' over-reliance on ready-made assessment materials. This limited knowledge about LAL is worrying because it can create conditions for inappropriate and wrong tests that might not assess the intended language outcomes, leading to test fairness and ethics issues. As a result, this study concludes that PD courses need to make this component explicit because it is the only condition that will lead to high-quality assessments for improving learners' language proficiency. (Giraldo, 2018; Stiggins, 2004).

Finally, the website analysis found similar results to Lam's (2015) analysis of official websites and interviews with undergraduate students and instructors about English Language Teaching Programs at five Hong Kong universities. Lam (2015) revealed a LAL theory-practice gap which was also found in my study. Therefore, it can be concluded that EFL teachers in Kazakhstan would probably have good knowledge about assessment in the updated curriculum

and new assessment terms such as criteria-based assessment, summative assessment, and formative assessment.

**Research Question 2: What are teachers' skills and knowledge about language assessment?**

Drawing on the modified list of Giraldo's (2018) descriptions for LAL dimensions, the survey results indicated the overall high level of teachers' self-perceptions about their abilities and knowledge (average mean 3.99 out of maximal five points scale). Namely, the study found four well-defined indicators of teachers' knowledge, skills, and principles in language assessment, including three positive and two negative ones:

First, the survey indicated that teachers are more confident in their abilities to conduct assessment processes ethically. It suggests that they learned about the importance and the functions of ethical assessment procedures from contact with IELTS exams which are used as a primary measurement for both teachers' and students' English language abilities in Kazakhstan. It is surprising because most literature points to the lack of focus on assessment principles in language testing textbooks (Brown & Bailey, 2008). In the Kazakh context, it could be explained by the consequential and powerful natures of large-scale tests such as IELTS and TOEFL are widely used as the leading indicator of language knowledge (Shohamy, 2001; Spolsky, 2008; Davies, 2008).

The second finding was teachers' high level of confidence in their instructional skills that emphasised teachers' self-perceived perceptions about their abilities to identify assessment purposes, design assessment tools, and prepare the tasks rubrics. It was also echoed in the document analysis, where the results suggested that the state-funded PD courses include the knowledge and skills related to selecting assessment tools and improving instructions based on assessment data results. Furthermore, the correlation analysis revealed that teachers who attend

PD interventions other than state-funded ones indicated higher confidence in the LAL skills component. This result showed that teachers who were actively engaging with additional professional development opportunities tend to be more prepared to improve their instructions and use technology in assessment practices.

The study found no statistically significant correlation between state-funded PD courses and teachers' LAL, which meant that the courses did not influence teachers' LAL development. Therefore, the teachers were not confident in their knowledge of language testing concepts concerning applied linguistics and second language acquisition. It was a disappointing finding because the Knowledge component includes teachers' understanding of communicative language testing, second language acquisition, and the development of appropriate assessment strategies and critical issues about validity and reliability (Inbar-Lourie, 2008; Davies, 2008). Therefore, the study concludes that EFL teachers' low LAL confidence can negatively affect classroom-based assessment practices because those with good knowledge of language assessment theories and principles can choose tests that best fit their classroom situations (Giraldo, 2019).

### **Research Question 3: What are teachers' beliefs and attitudes about language assessment?**

First, the study found that the participants' self-perceived LAL perceptions were average, which concurs with other large-scale surveys that found teachers' self-perceived LAL was regrettably low (DeLuca & Lam, 2014; Vogt & Tsagari, 2014; Mertler, 2009; Lam, 2015; Xu & Brown, 2017). Unfortunately, it holds negative implications for EFL classrooms because an average LAL can limit pedagogy, narrow the scope of language learning, and suggest that the teachers would probably be more comfortable designing and reporting on summative assessment tasks for grade completion (Chan & Evans, 2014; López & Bernal, 2009).

Secondly, the study found that teachers demonstrated low confidence in developing assessments, indicative of pedagogical and technical knowledge gaps, despite attending assessment PD courses. For this reason, the data showed an over-reliance on assessment guides, which highlighted the participants' limited capacity to draw conclusions about the academic achievements of their EFL students (Alkharusi et al., 2011; Popham, 2009). As a result, the study concludes that the participants' responses did not demonstrate 1) “an increased awareness of and appreciation for assessment as a tool for guiding and improving language instruction” (Montee et al., 2013, p. 23) or 2) nor did they demonstrate assessment as “a process in which the teacher gathers relevant information about the student’s weaknesses and strengths in the learning process to make decisions about the instruction and students’ learning” (Restrepo-Bolívar, 2020, p. 45).

The third finding illustrated that teachers’ beliefs and values foregrounded summative assessment and institutional accountability. It was illuminating that their professional identities did not make explicit the relationship between models of language pedagogy and assessment design, their educational measurement designs, and the impact of language testing on their classroom-based practices (Davies, 2008). For example, their LAL beliefs and professional identities did not include assessing core language skills such as listening, speaking, reading, and writing. Therefore, the results demonstrated that teachers might find it challenging to articulate and shift assessment practices from grammar and vocabulary to optimally assessing communicative competence.

Finally, the study found a positive teachers’ stance toward creating conditions for the co-construction of assessment knowledge collaboration and mentoring rather than attending more PD sessions. Another dominant finding was the theoretical nature of assessment materials and initiatives, which suggested they were probably not exposed to practical, hands-on training

(Fulcher, 2012; Giraldo & Murcia, 2018). This finding emphasised that if teachers study LAL knowledge and principles, they should do so within a practice-based framework (Kleinsasser, 2005; Restrepo-Bolívar, 2020). In sum, the development of language teachers' LAL should happen as a bottom-up intention where all the essential aspects of teachers' needs and challenges are taken into account (Fulcher, 2012; Scarino, 2013; Giraldo, 2018).

**Main Research Question: To what extent do teachers' language assessment literacy (LAL) reflect the underpinnings of language teaching reform?**

The study found that the state-funded PD organisations' course objectives are working towards improving EFL teachers' professional skills, including assessment knowledge, which was visible in the information on their websites. However, the focus was on developing general pedagogical knowledge and introducing criteria-based assessment and summative and formative assessment strategies. Therefore, the study found that significant LAL components were not covered in PD courses, revealing a LAL theory-practice gap (Lam, 2015). Based on the assessment PD courses, the study concludes that EFL teachers in Kazakhstan would probably have some assessment knowledge to navigate the updated curriculum. As a result, the study concludes that PD courses need to make the knowledge component that consists of the relationship between applied linguistics and language testing because it is the only condition that will lead to high-quality assessments for improving learners' language proficiency. (Giraldo, 2018; Stiggins, 2004).

The quantitative study revealed three well-defined trends. First, the results showed that the teacher's completion of other PD interventions related to teachers' LAL skills ( $r = -0.22$  and  $-0.15$ ,  $p < 0.05$ ). Second, and in contrast, there was no statistically significant correlation between state-funded PD courses and teachers' LAL, meaning that completing the courses did not develop

teachers' LAL. Third, within the Skills component, the completion of PD intervention was related to teachers' Instructional and Design skills ( $r = -0.19$  and  $-0.21$ ,  $p < 0.05$ ). According to their self-reported responses, each correlation was negative, meaning those who did not attend PD interventions seemed to have lower abilities. As a result, the state-funded PD did not develop teachers' LAL, which was confirmed by the statistically significant weak negative correlation between the items related to teachers' completions of PDs and their self-indicated language assessment skills and knowledge. Consequently, the study found a correlation between the LAL gap in the PD courses and participants' self-perceived LAL knowledge.

Finally, the qualitative study results indicated that 1) the participants' self-perceived LAL perceptions were average, 2) gaps in teachers' pedagogical and technical knowledge resulting in low confidence in developing assessments, 3) the teachers developed LAL professional identities related to summative assessment and institutional accountability and 4) the participants held a positive stance toward collaboration about the co-construction of assessment knowledge rather than PD sessions. It is worth mentioning that the results mentioned above are in correlation with the teacher survey findings, where most teachers reported their low preparedness to employ descriptive statistics, knowledge of applied linguistics and language testing theory and principles in their assessment practices.

## **Conclusion**

This chapter discussed the findings presented in Chapter 4 in light of existing theoretical and empirical studies related to the current research topic and Giraldo's (2018) descriptions for LAL dimensions, which were employed as the conceptual framework of this study. Thus, this chapter foregrounded the main findings on EFL teachers' self-perceived LAL, the nature of state-funded EFL assessment. In addition, this chapter illustrated teachers' perceptions and beliefs

about language assessment, and their suggestions about improving teachers' awareness of LAL in local and national contexts.

## Chapter 6: Conclusion

The previous chapter focused on presenting findings and discussions from qualitative and quantitative data gathered from document analysis, teacher survey and teacher interviews. The document analysis included three English as a foreign language (EFL) teacher training course descriptions taken from the websites of two state-funded PD organisations. The teacher survey was completed by 119 participants, where EFL teachers voluntarily answered 5-point Likert-type survey questions that used a modified list of Giraldo's (2018) descriptions for LAL dimensions. Finally, the semi-structured teacher interviews were conducted with three participants, representing in-service EFL teachers from urban and rural schools, and a pre-service EFL teacher. For the analysis of all data, the study utilised Giraldo's (2018) descriptions for LAL dimensions which was employed as the conceptual framework of this study. As stated in the Introduction, the purpose of the study was to explore the EFL teachers' LAL level and shed light on how it is impacted by state-funded assessment PD courses. The central question of the study was:

To what extent do teachers' language assessment literacy (LAL) reflect the underpinnings of language teaching reform?

The following sub-question guided the main question:

1. What is the content of language assessment PD training in Kazakhstan?
2. What are teachers' skills and knowledge about language assessment?
3. What are teachers' beliefs and attitudes about language assessment?

In the concluding chapter, I foreground and summarise the findings of the above research questions. The study's main conclusions will be outlined: 1) 1) The content of language assessment PD training in Kazakhstan, 2) Teachers' skills and knowledge about language

assessment, and 3) The relationship between teachers' language assessment literacy (LAL) and language teaching reform. Furthermore, this chapter presents the implications of the results, the study's limitations and provides recommendations for further research.

### **Main Conclusions of the Study**

#### ***The content of language assessment PD training in Kazakhstan***

The study revealed that state-funded PD organisations work successfully on improving EFL teachers' assessment skills and knowledge. However, it was also found that the PD courses they provide are primarily focused on sharing assessment knowledge related to the recent school curriculum update, such as the implementation of criteria-based assessment, formative and summative assessment, and the materials and practices for developing and analysing various assessment tools. What is not present in the courses was the significant component of LAL, the knowledge that includes applied linguistics such as theories of second language acquisition, critical language testing, and multilingualism. The same results were also echoed (a) in the online teacher survey, where there was no statistically significant correlation between state-funded PD courses and teachers' LAL, meaning that completing the courses did not develop teachers' LAL, and (b) in the teacher interviews, where the respondents' self-perceived LAL perceptions were average. The results indicate the needs of in-service EFL teachers that are not covered in the PD curriculum.

#### ***Teachers' skills and knowledge about language assessment***

The study had led the researcher to conclude that participants' self-perceived LAL was lower than average. Although most survey participants were optimistic about their skills required for language assessment processes, they had low responses for the items of knowledge component, indicating that they were not confident in their knowledge of language testing

concepts concerning applied linguistics and second language acquisition. It was also echoed in the results of the interviews, where teachers' assessment beliefs and values foregrounded summative assessment and institutional accountability; and their professional identities did not make explicit the relationship between models of language pedagogy and assessment design, their educational measurement designs, and the impact of language testing on their classroom-based practices. Taken together, it suggested that EFL teachers might have gaps in theoretical and technical knowledge resulting in low confidence in developing and managing language assessments, which can negatively affect their classroom-based assessment practices.

***The relationship between teachers' language assessment literacy (LAL) and language teaching reform***

Unfortunately, despite the growing interest in the high-quality English language teaching in the country, there is a tendency to neglect the needs of EFL teachers, in which the top-down designed professional development programs fail to provide the teachers with the assessment knowledge and skills. It obviously leads to the decrease in language teaching quality in public schools, which in turn challenges the implementation of the current language-in-education reforms in Kazakhstan. The problem might be resolved by focusing on improving EFL teachers' language assessment literacy (LAL), which can include all the mentioned gaps in the knowledge of the EFL teachers. Finally, it is essential to treat LAL as a significant problem and to contextualise it with the local educational system and, most importantly, to teachers' workforce, levels of teachers and students, meaning that the LAL of teachers assessing young learners might differ in those who work with advanced levels or higher grades of school (Tsagari, 2021).

### **Limitations of the Study**

The study has several limitations. Firstly, the study results cannot be generalised to the whole population of EFL teachers in the country as it was small scale research. Secondly, one potential limitation of the interview analysis is based on the languages used for conducting the interviews. The interviews used Kazakh and Russian languages as they were the participants' first languages, and for the analysis, the transcripts were translated into English. It might result in the quotes which were sometimes a bit complicated. To minimise its impact and not lose any vital information, the interviews were coded in their original languages first and then translated all the quotes into English. Thirdly, the series of mass protests in Kazakhstan in January 2022 led to the nationwide internet blackout for almost ten days, which highly impacted the possible numbers of the respondents for the teacher survey. Moreover, the website analysis of the state-funded PD organisations could not provide sufficient information about the nature and context of their interventions for EFL teachers. Thus, some of the results of the document analysis step could be biased.

Lastly, all collected data was based on the website analysis and the different perceptions of EFL teachers throughout the country to increase the chances of getting more accurate information about the current state of the EFL teachers' LAL and its relation to the aims and processes of state-funded PD organisation. Triangulation allowed the collecting of valuable data. However, at the same time, most of the data were based on the assumptions and perceptions of the respondents and interviewees and finally depended on the personal interpretation of the investigator.

### **Implications and Recommendations**

The findings of this study have various implications and recommendations for the development of teacher LAL in Kazakhstan. Although there is a growing interest in LAL studies worldwide, there is a gap in language teachers' LAL studies in Kazakhstan. Therefore, the topic of EFL teachers' LAL in Kazakhstan and the impact of state-funded professional development organisations need further research.

#### ***Practice***

Based on the study's findings, an assessment PD curriculum to foster EFL teachers' LAL should focus on the knowledge dimension. Although the high-level knowledge of theory is the least important aspect for classroom teachers (Taylor, 2013), the presence of introductory knowledge about applied linguistics, critical language testing, and multilingualism in language teachers' assessment PD courses in Kazakhstan is a must. In addition, taking into account that a PD curriculum has a high role in developing teachers' assessment and LAL, and LAL enhancement is collaborative work, policymakers in Kazakhstan need to shift from the top-down policy in developing courses for the educators and start the evaluation of teachers' needs and include the knowledge on LAL to every EFL teachers' course on assessment. In addition, the same implication is appropriate to the stakeholders from EFL teaching institutions.

#### ***Future Research***

Future work should concentrate on enhancing the quality of assessment PD courses for in-service and pre-service EFL teachers. Therefore, it is suggested to investigate and compare the LAL needs of EFL teachers of different levels (primary, secondary, high schools, and private tutoring) in Kazakhstan. Future needs would be to prepare contextualised LAL profiles for different types of EFL teachers or to establish professional standards and guidelines for EFL

teachers' LAL. In addition, the same investigation can be done with teacher trainers and, most importantly, on the professional language testers who deal with the development and coordination of the high-stake English language tests in the country's education system (e.g., Unified National Testing). Another possible area is to explore how EFL teachers in the country develop their language assessment knowledge and skills via other sources than state-funded courses or how they choose tests for their classroom situations and examine their impact on teachers' perceptions about their own LAL.

Finally, it is again worth mentioning that LAL is not a theoretical model that can be used as a unified tool to measure language teachers' assessment abilities in different contexts (Sun, 2022). And taking into account that LAL is essential in improving the English language teaching, for Kazakhstan, it is a significant issue and needs to be contextualised within the demands of the local educational system, considering various factors such as the levels of students, teachers' workforce, the difference between teachers from rural and urban schools, or teachers assessing young learners and students of advanced levels (Tzagari, 2021).

### References

- Akhmedina, A. (2017). Challenges of implementing formative assessment at Nazarbayev Intellectual School. [Master's thesis, Nazarbayev University]. Nazarbayev University Repository. <http://nur.nu.edu.kz/handle/123456789/2555>
- Alkharusi, H., Kazem, A. M., & Al-Musawai, A. (2011). Knowledge, skills, and attitudes of preservice and inservice teachers in educational measurement. *Asia-Pacific Journal of Teacher Education*, 39(2), 113-123.
- Alkharusi, H. (2011). Psychometric properties of the Teacher Assessment Literacy Questionnaire for preservice teachers in Oman. In *Procedia - Social and Behavioral Sciences* (Vol. 29, pp. 1614–1624). <https://doi.org/10.1016/j.sbspro.2011.11.404>
- Bachman, L. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bachman, L. F., & Palmer, A. S. (2010). *Language testing in practice*. Oxford: Oxford University Press.
- Bachman, L. F. and Palmer, A. (2010). *Language assessment in practice: Developing language assessments and justifying their use in the real world*. Oxford, UK: Oxford University Press.
- Black, P., & Wiliam, D. (2010). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 92(1), 81–90.  
<https://doi.org/10.1177/003172171009200119>
- Bondarenko, M. (2019). Influences of criteria-based assessment on English language teachers' teaching practices: a case of one mainstream secondary school in Kazakhstan. [Master's

- thesis, Nazarbayev University]. Nazarbayev University Repository.  
<http://nur.nu.edu.kz/handle/123456789/4339>
- Bonett, D. G., & Wright, T. A. (2015). Cronbach's alpha reliability: Interval estimation, hypothesis testing, and sample size planning. *Journal of Organizational Behavior*, 36(1), 3–15. <https://doi.org/10.1002/job.1960>
- Borg, S. (2015). *Teacher cognition and language education: Research and practice*. Bloomsbury Publishing.
- Brookhart, S. M. (2011). Educational Assessment Knowledge and Skills for Teachers. *Educational Measurement: Issues and Practice*, 30(1), 3–12.  
<https://doi.org/10.1111/j.1745-3992.2010.00195.x>
- Brown, D. J., & Bailey, K. M. (2008). Language testing courses: What are they in 2007? *Language Testing*, 25(3), 349–383. <https://doi.org/10.1177/0265532208090157>
- Bryman, A. (2016). *Social research methods*. Oxford university press.
- Campbell, C., Murphy, J. A., & Holt, J. K. (2002). Psychometric analysis of an assessment literacy instrument: Applicability to pre-service teachers. Paper presented at the annual meeting of the Mid-Western Educational Research Association, Columbus, OH
- Canagarajah, S. (2006). Changing communicative needs, revised assessment objectives: Testing English as an international language. *Language Assessment Quarterly*, 3(3), 229–242.
- Cenoz, J., & Gorter, D. (2020). Pedagogical translanguaging: An introduction. *System*, 92.  
<https://doi.org/10.1016/j.system.2020.102269>
- Chan, J. Y. H., & Evans, S. (2011). Choosing an appropriate pronunciation model for the ELT classroom: A Hong Kong perspective. *Journal of Asia TEFL*, 8(4), 1–24.

- Creswell, J. (2014). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research: International edition* (4th ed.). Boston: Pearson.
- Creswell, J. W., & Clark, V. L. P. (2018). Designing and Conducting Mixed Methods Research. *Organizational Research Methods*, 12(4), 801–804.
- Dahm, R., & De Angelis, G. (2018). The role of mother tongue literacy in language learning and mathematical learning: is there a multilingual benefit for both? *International Journal of Multilingualism*, 15(2), 194–213. <https://doi.org/10.1080/14790718.2017.1359275>
- Davies, A. (2004). Introduction: Language Testing and the Golden Rule. *Language Assessment Quarterly: An International Journal*, 1(2&3), 97–107.  
[https://doi.org/10.1207/s15434311laq12&3\\_2](https://doi.org/10.1207/s15434311laq12&3_2)
- Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing*, 25(3), 327–347 <https://doi.org/10.1177/0265532208090156>.
- De Backer, F., Slembrouck, S., & Van Avermaet, P. (2019). Language Passports: The Multilayered Linguistic Repertoires of Multilingual Pupils in Flanders, Belgium. In *Handbook of the Changing World Language Map* (pp. 1–16). Springer International Publishing. [https://doi.org/10.1007/978-3-319-73400-2\\_161-1](https://doi.org/10.1007/978-3-319-73400-2_161-1)
- 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>
- DeLuca, C., & Lam, C. Y. (2014). Preparing Teachers for Assessment within Diverse Classrooms: An Analysis of Teacher Candidates' Conceptualizations. *Teacher Education Quarterly*, 41(3), 3-24.

- Dendrinios, B. (2018). Multilingualism language policy in the eu today: A paradigm shift in language education. *Training, Language and Culture*, 2(3), 9–28.  
<https://doi.org/10.29366/2018tlc.2.3.1>
- DuFour, R., DuFour, R., Loertscher, D. V., & Eaker, R. (2010). Learning by Doing: A Handbook for Professional Learning Communities at Work (Second Edition). *Solution Tree Press*, 75.
- 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., & Murcia, D. (2018). Language Assessment Literacy for Pre-service Teachers: Course Expectations from Different Stakeholders. *GiST Education and Learning Research Journal*, (16), 56–77. <https://doi.org/10.26817/16925777.425>
- Giraldo, F. (2018). Language assessment literacy: implications for language teachers. *Profile: Issues in Teachers' Professional Development*, 20(1), 179–195  
<https://doi.org/10.15446/profile.v20n1.62089>
- Harding, L., & Kremmel, B. (2016). Teacher assessment literacy and professional development. In *Handbook of Second Language Assessment* (pp. 413–427). De Gruyter.  
<https://doi.org/10.1515/9781614513827-027>
- Harlen, W. (2012). The role of assessment in developing motivation for learning. In *Assessment and Learning* (pp. 171–184). SAGE Publications Inc.  
<https://doi.org/10.4135/9781446250808.n11>
- Herrera, L., & Macías, D. F. (2015). A call for language assessment literacy in the education and development of teachers of English as a foreign language. *Colombian Applied Linguistics Journal*, 17(2), 302–312. <https://doi.org/10.14483/udistrital.jour.calj.2015.2.a09>

Heugh, K., Prinsloo, C., Makgamatha, M., Diedericks, G., & Winnaar, L. (2017).

Multilingualism(s) and system-wide assessment: a southern perspective. *Language and Education*, 31(3), 197–216. <https://doi.org/10.1080/09500782.2016.1261894>

Hidri, S. (2020). Language assessment literacy. In *Perspectives on Language Assessment Literacy* (pp. 3–12). Routledge. <https://doi.org/10.4324/9781003016083-1>

Ilic, P., & Warrington, S. D. (2007). Pre-testing the Annual Activities Survey for Language Teachers (AASLT). *The Asian EFL Journal Quarterly* 9(4), 167.

Inbar-Lourie, O., Scarino, A., Malone, M. E., Jeong, H., O'Loughlin, K., Pill, J., & Harding, L. (2013). The special issue on language assessment literacy. *Language Testing*, 30(3), 301–307.

Inbar-Lourie, O. (2008). Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing*, 25(3), 385–402.  
<https://doi.org/10.1177/0265532208090158>

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

Karabassova, L. (2020). Understanding Trilingual Education Reform in Kazakhstan: Why Is It Stalled? (pp. 37–51). [https://doi.org/10.1007/978-3-030-50127-3\\_3](https://doi.org/10.1007/978-3-030-50127-3_3)

Karabassova, L. (2021). English-medium education reform in Kazakhstan: comparative study of educational change across two contexts in one country. *Current Issues in Language Planning*, 22(5), 553–573. <https://doi.org/10.1080/14664208.2021.1884436>

Klein, E. D., & van Ackeren, I. (2011). Challenges and problems for research in the field of statewide exams. A stock taking of differing procedures and standardization levels.

*Studies in Educational Evaluation*, 37(4), 180–188.

<https://doi.org/10.1016/j.stueduc.2012.01.002>

Kleinsasser, R. C. (2005). Transforming a postgraduate level assessment course: A second language teacher educator's narrative. *Prospect*, 20, 77–102

Kramsch, C. (1993). *Context and culture in language teaching*. Oxford: Oxford University Press.

Kramsch, C. (Ed.) (2014). Teaching foreign languages in the era of globalization. *The Modern Language Journal*, 98(Special issue 1), 296.

Kunnan, A. J. (2004). Test fairness. In M. Milanovic, C. Weir, & S. Bolton (Eds.), *Europe language testing in a global context: Selected papers from the ALTE conference in Barcelona* (pp. 27–48). Cambridge: Cambridge University Press.

Lam, R. (2015). Language assessment training in Hong Kong: Implications for language assessment literacy. *Language Testing*, 32(2), 169–197.

López, A., & Bernal, R. (2009). Language testing in Colombia: A call for more teacher education and teacher training in language assessment. *Profile: Issues in Teachers' Professional Development*, 11(2), 55-70.

Lopez, A. A., Turkan, S., & Guzman-Orth, D. (2017). Assessing Multilingual Competence. In *Language Testing and Assessment* (pp. 91–102). Springer International Publishing.

[https://doi.org/10.1007/978-3-319-02261-1\\_6](https://doi.org/10.1007/978-3-319-02261-1_6)

Malone, M. E. (2013). The essentials of assessment literacy: Contrasts between testers and users. *Language Testing*, 30(3), 329–344.

McMillan, J. H., & Nash, S. (2000). Teacher classroom assessment and grading practices decision making. *Paper Presented at the Annual Meeting of the National Council on Measurement in Education*, 39.

- McNamara, T., & Roever, C. (2006). *Language testing: The social dimension*. Oxford: Blackwell Publishing.
- McNamara, T., & Shohamy, E. (2008). Language tests and human rights. *International Journal of Applied Linguistics*, 18(1), 89–95.
- Meier, G. S. (2017). The multilingual turn as a critical movement in education: Assumptions, challenges and a need for reflection. *Applied Linguistics Review*, 8(1), 131–161.  
<https://doi.org/10.1515/applirev-2016-2010>
- Menken, K. (2008). *English Learners Left Behind Standardized Testing as Language Policy*. *English Learners Left Behind: Standardized Testing as Language Policy* (pp. 1–207). Channel View Publications. <https://doi.org/10.1080/13670050802499662>
- 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>
- Ministry of Education and Science of the Republic of Kazakhstan [MoES]. (2015). *State Program of Education Development for 2016-2019*.  
[http://online.zakon.kz/Document/?doc\\_id=31024348#pos=0;920](http://online.zakon.kz/Document/?doc_id=31024348#pos=0;920)
- Montee, M., Bach, A., Donovan, A., & Thompson, L. (2013). LCTL teachers' assessment knowledge and practices: An exploratory study. *Journal of the National Council of Less Commonly Taught Languages*, 13, 1-31.
- Moss, C. M. (2013). Research on classroom summative assessment. In *SAGE Handbook of Research on Classroom Assessment* (pp. 235–256). SAGE Publications Inc.  
<https://doi.org/10.4135/9781452218649.n14>

- OECD (Organisation for Economic Co-operation and Development). (2014). Reviews of national policies for education: secondary education in Kazakhstan. Paris: OECD Publishing.
- O'Sullivan, B. (2002). Learner acquaintanceship and oral proficiency test pair-task performance. *Language Testing*, 19(3), 277–295.
- Pennycook, A. (2001). *Critical applied linguistics: A critical introduction*. Mahwah: Lawrence Erlbaum.
- Plake, B. S., & Impara, J. C. (1992). *Teacher competencies questionnaire description*. Lincoln, NE: University of Nebraska.
- Plake, B. S., Impara, J. C., & Fager, J. J. (1993). Assessment competencies of teachers: A national survey. *Educational Measurement: Issues and Practice*, 12(4), 10-12.
- Polesel, J., Rice, S., & Dulfer, N. (2014). The impact of high-stakes testing on curriculum and pedagogy: A teacher perspective from Australia. *Journal of Education Policy*, 29(5), 640–657. <https://doi.org/10.1080/02680939.2013.865082>
- Popham, W. J. (2004). Why assessment illiteracy is professional suicide. *Educational Leadership*, 62(1), 82–83.
- Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory into Practice*, 48(1), 4–11. <https://doi.org/10.1080/00405840802577536>
- Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory into Practice*, 48(1), 4–11. <https://doi.org/10.1080/00405840802577536>
- Razmjoo, S. A. (2007). High schools or private institutes textbooks? Which fulfill communicative language teaching principles in the Iranian context. *Asian EFL journal*, 9(4), 126-140.

- Restrepo-Bolívar, E. M. (2020). Monitoring Preservice Teachers' Language Assessment Literacy Development through Journal Writing. *Malaysian Journal of Elt Research*, 17(1), 38–52.
- Sachtleben, A., & Crezee, I. (2015). Digital innovation serving interpreter education in New Zealand. In *Interpreter Education in the Digital Age: Innovation, Access, and Change* (pp. 18–38). Gallaudet University Press.
- Scarino, A. (2010). Assessing intercultural capability in learning languages: A renewed understanding of language, culture, learning and the nature of assessment. *Modern Language Journal*, 94(2), 324–329. doi:10.1111/j.1540-4781.2010.01026.x.
- 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>
- Shamshidinova, K., Ayubayeva, N., & Bridges, D. (2014). Implementing radical change: Nazarbayev intellectual schools as agents of change. In D. Bridges (Ed.), *Educational reform and internationalisation: The case of school reform in Kazakhstan* (pp. 71–82). Cambridge, UK: Cambridge University Press.
- Shohamy, E. (1992). *The power of tests: a study of the impact of language tests on teaching and learning*. Paper presented at the Language Testing Research Colloquium, Vancouver.
- Shohamy, E. (2001). *The power of tests: A critical perspective on the uses of language tests*. London: Longman/Pearson Education.
- Shohamy, E. (2006). *Language policy: Hidden agendas and new approaches*. New York: Routledge.
- Shohamy, E. (2007). Language tests as language policy tools. *Assessment in Education*, 14(1), 117-130.

- Shohamy, E. (2011). Assessing multilingual competencies: Adopting construct valid assessment policies. *Modern Language Journal*, 95(3), 418–429. <https://doi.org/10.1111/j.1540-4781.2011.01210.x>
- Shohamy, E. (2021). Critical language testing. In *The Power of Tests* (pp. 157–160). Routledge. <https://doi.org/10.4324/9781315837970-30>
- Shohamy, E. and Menken, K. (2015). Language Assessment. In *The Handbook of Bilingual and Multilingual Education* (eds W.E. Wright, S. Boun and O. García). <https://doi.org/10.1002/9781118533406.ch15>
- Spolsky, B. (1977). Language testing: An art or science. In *Proceedings of the fourth international congress of applied linguistics* (pp. 7–28). Stuttgart: Hochschulverlag.
- Spolsky, B. (2008). Language testing at 25: Maturity and responsibility. *Language Testing*, 25(3), 297–305.
- Spolsky, B. (2017). History of Language Testing. In *Language Testing and Assessment* (pp. 375–384). Springer International Publishing. [https://doi.org/10.1007/978-3-319-02261-1\\_32](https://doi.org/10.1007/978-3-319-02261-1_32)
- Stavans, A., & Hoffmann, C. (2015). *Multilingualism*. Cambridge: Cambridge University Press.
- Stiggins, R. (2004). New assessment beliefs for a new school mission. *Phi Delta Kappan*, 86, 22–27 <https://doi.org/10.1177/003172170408600106>
- Stiggins, R. J. (1991). Assessment literacy. *The Phi Delta Kappan*, 72(7), 534–539.
- Stobart, G. (2008). *Testing times: The uses and abuses of assessment*. *Testing Times: The Uses and Abuses of Assessment* (pp. 1–218). Routledge. <https://doi.org/10.4324/9780203930502>

- Sun, Y. (2022). A Review and Discussion of In-service EFL Teachers' Language Assessment Literacy Level in Junior High School in China. *Studies in Applied Linguistics and TESOL*, 21(2). <https://doi.org/10.52214/salt.v21i2.9060>
- Tahereen, T. (2014). Testing based language teaching in Bangladesh: Does it promote or impede learning. *TESOL International Journal*, 9(1), 148-178.
- Tastanbekova, K. (2020). Professional prestige, status and esteem of teaching in Kazakhstan: Temporal, regional and gender analysis of payroll data. *Journal of Eastern European and Central Asian Research*, 7(2), 175–190.
- 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.
- Tsagari, D. (2021). Gauging the assessment literacy levels of English language teachers in Norway. *European Journal of Applied Linguistics and TEFL*, 10(1), 161–191.
- Vélez-Rendón, G. (2006). From student to teacher: A successful transition. *Foreign Language Annals*, 39(2), 320-333.
- Vogt, K., & Tsagari, 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>
- Wyatt-Smith, C., Klenowski, V., & Gunn, S. (2010). The centrality of teachers' judgement practice in assessment: A study of standards in moderation. *Assessment in Education: Principles, Policy and Practice*, 17(1), 59–75.

Xu, Y., & Brown, G. T. L. (2017). University English teacher assessment literacy: A survey-test report from China. *Papers in Language Testing and Assessment*, 6(1), 133–158.

## Appendix A

### Five-Point Likert-Type Scale Questionnaire

#### Step 1

Participants' info, teaching experience and education, PD experience

<b>Participants' Info</b>	Gender: Age: Type of school you work at: Urban Rural Grades you work with (more than one option): Primary school (K-4) Secondary school (5-9) High school (10-11, and 12) Last teacher category: Teacher, Teacher-Moderator, Teacher-Expert, Teacher-Researcher Teacher-Master
<b>Teaching Experience &amp; Education</b>	Length of time teaching English, in years (indicate the number): Choose your degree or qualification (more than one option): PhD (graduated or near completion) MA BA Other: Certificate (CELTA, DELTA, TEFLA etc.)
<b>Participation in PD Interventions</b>	Have you taken any state-funded professional development (PD) training on the updated curriculum (Orleu, CoE)? Yes/No. If yes, indicate the year and the organisation: Have you taken any PD intervention on assessment (seminar, workshop, training, webinar etc.)? Yes/No. If yes, indicate the year and the organisation:

**Step 2**

The current state of EFL teachers' LAL

Please indicate how skilled you are in using the assessment issue described by each of the statements below by ticking one of the choices next to each statement according to the following five-point scale:

SA= Strongly agree,  
A=Agreed,  
NS = Not sure,  
D = Disagree,  
SD = Strongly disagree.

<b>Skills</b>	
<p><b><i>Instructional skills</i></b> I have ability to:</p>	<ol style="list-style-type: none"> <li>1. plan, implement, monitor, record, and report student language development.</li> <li>2. improve instruction based on assessment results and feedback.</li> <li>3. utilize alternative means for assessment; for example, portfolios.</li> <li>4. communicate norm- and criterion-referenced test results to a variety of audiences: students, parents, school directors etc.</li> </ol>
<p><b><i>Design skills for language assessments</i></b> I have ability to:</p>	<ol style="list-style-type: none"> <li>1. clearly identify and state the purpose for language assessment.</li> <li>2. write selected-response items such as multiple-choice, true-false, and matching.</li> <li>3. design rubrics for alternative assessments such as portfolios and peer-assessment.</li> </ol>
<p><b><i>Technological skills</i></b> I have ability to:</p>	<ol style="list-style-type: none"> <li>1. run operations on Excel; for example, descriptive statistics and reliability correlations.</li> <li>2. use internet resources such as online tutorials and adapt contents for their particular language assessment needs.</li> </ol>
<b>Knowledge</b>	
<p><b><i>Awareness of applied linguistics</i></b> I find it easy to:</p>	<ol style="list-style-type: none"> <li>1. compare approaches for language teaching and assessment; e.g., communicative language testing; task-based assessment.</li> <li>2. explain the major issues in applied linguistics; e.g., bilingualism, language policy and planning, pragmatics, sociolinguistics, etc.</li> <li>3. analyse trends in second language acquisition and their</li> </ol>

	impact on language assessment; e.g., motivation, cross-linguistic influence, learner strategies.
<i>Awareness of theory and concepts</i> I find it easy to:	<ol style="list-style-type: none"> <li>1. illustrate the history of language testing and assessment, and its impact on current practices and society.</li> <li>2. evaluate the kind of washback that assessments can have on learning, teaching, curricula, and institutions.</li> </ol>
<b>Principles</b>	
<i>Awareness of and actions towards critical issues in language assessment</i> I find it easy to:	<ol style="list-style-type: none"> <li>1. use tests, test processes, and test scores ethically.</li> <li>2. provide assessment practices that are fair and non-discriminatory.</li> <li>3. implement democratic language assessment practices, by giving students opportunities to share their voices about assessment.</li> </ol>

\*The questions are modified Giraldo's (2018) eight dimensions of LAL for language teachers.

## Appendix B

### Interview Questions

Categories	Interview Questions
<b>Warm-up Questions</b>	<ol style="list-style-type: none"> <li>1. How long have you been teaching English at school?</li> <li>2. Were you involved in state-funded teacher training from organisations like Orleu or Center of Excellence?</li> </ol> <p><b>Probes:</b> What types of interventions did you participate in? (Training, seminar, workshop etc.)</p>
<b>What are teachers' attitudes and beliefs about language assessment in classrooms?</b>	<ol style="list-style-type: none"> <li>1. Did you work at school before the updated curriculum reform was introduced?               <ol style="list-style-type: none"> <li>0. What assessment models did you use at that time?                   <ol style="list-style-type: none"> <li>1. Do you see any differences or similarities between the traditional and updated assessment models?</li> </ol> </li> </ol> </li> <li>2. What do you think about the teachers' assessment of literacy?               <ol style="list-style-type: none"> <li>0. How does it affect their professionalism?                   <ol style="list-style-type: none"> <li>1. How does it affect the language teachers' professionalism?</li> </ol> </li> </ol> </li> <li>3. What is your self-perceived level of language assessment literacy (LAL)? What do you think are the reasons for your answer?</li> </ol> <p><b>Probes:</b> What is the most difficult aspect of LAL in your classroom practice? Why is this?</p>
<b>What are the teachers' suggestions about language assessment in classrooms?</b>	<ol style="list-style-type: none"> <li>1. What do you think should be done to improve the teachers' LAL in the school you work in? What about at the national level?</li> </ol> <p><b>Probes:</b> What kind of PD interventions would be useful to you?</p>

## Appendix C

### Informed Consent Form

#### **An investigation of English as a Foreign Language (EFL) teachers' assessment literacy in Kazakhstan**

**DESCRIPTION:** You are invited to participate in a research study that explores your attitudes and beliefs about language assessment literacy (LAL) and your experiences in professional development (courses you have participated in, seminars or other interventions on assessment methodology). You will be asked to participate in a face-to-face interview or online interview, depending on the COVID 19 restrictions in Kazakhstan. The interviews will be informal and will be conducted in the English, Russian or Kazakh languages depending on your preference. If you express your agreement, the interview will be tape-recorded. In addition, you will be requested to complete a questionnaire about your knowledge on LAL (competence and assessment practice frequency). Finally, you will answer questions related to your attitude towards the updated assessment methodology, your professional development experience and your competence in LAL. Your name, the name of your organisation and other personal information will be anonymous during all stages of the study, including in documents, electronic files, and the dissertation itself. All the collected data, recordings, and study-related documents, including consent forms, will be saved on a personal cloud-based storage protected with strong password which is only accessible to the researcher.

**TIME INVOLVEMENT:** Your participation for each instrument will take between 20-30 minutes.

**RISKS AND BENEFITS:** The risks for participants associated with this study are very minimal. To reduce any potential risks arising from the study, participants' and their institution's identities will be hidden; thus, their names will be changed in the documents and the data stored in a secured place. No information from the interviews will be shared with their colleague teachers or administration. The interview time and place will be negotiated with every participant and will not interfere with their class schedules, work duties and important office hours. Therefore, participants will not lose their work time, and the study will not interrupt their regular work schedule.

The benefits that may reasonably be expected from this study are the participants having the opportunity to share their perspectives that are unique in Kazakhstani research and receiving new insights about the EFL teachers' state of LAL, their attitudes and beliefs about the changed

assessment methodology and whether the knowledge and practices of teachers corresponds to the aims and content of state-funded PD courses.

**PARTICIPANT'S RIGHTS:** If you have read this form and have decided to participate in this project, please understand your participation is voluntary. As such, you have the right to withdraw your consent or discontinue participation at any time without penalty. The alternative is not to participate and the right to refuse to answer particular questions. The results of this research study may be presented at scientific or professional meetings or published in scientific journals.

**CONTACT INFORMATION:**

*Questions:* If you have any questions, concerns or complaints about this research, its procedures, risks and benefits, please contact the Research Project Supervisor for this student work, Assistant Professor Michelle Bedeker, at [michelle.bedeker@nu.edu.kz](mailto:michelle.bedeker@nu.edu.kz)

*Independent Contact:* If you are not satisfied with how this study is being conducted, or if you have any concerns, complaints, or general questions about the research or your rights as a participant, please contact the NUGSE Research Committee to speak to someone independent of the research team at +7 7172 709359. You can also write an email to the NUGSE Research Committee at [gse\\_researchcommittee@nu.edu.kz](mailto:gse_researchcommittee@nu.edu.kz)

Please sign this consent form if you agree to participate in this study.

- I have carefully read the information provided;
- I have been given full information regarding the purpose and procedures of the study;
- I understand how the data collected will be used, and that any confidential information will be seen only by the researchers and will not be revealed to anyone else;
- I understand that I am free to withdraw from the study at any time without giving a reason;
- With full knowledge of all previous information, I agree, of my own free will, to participate in this study.

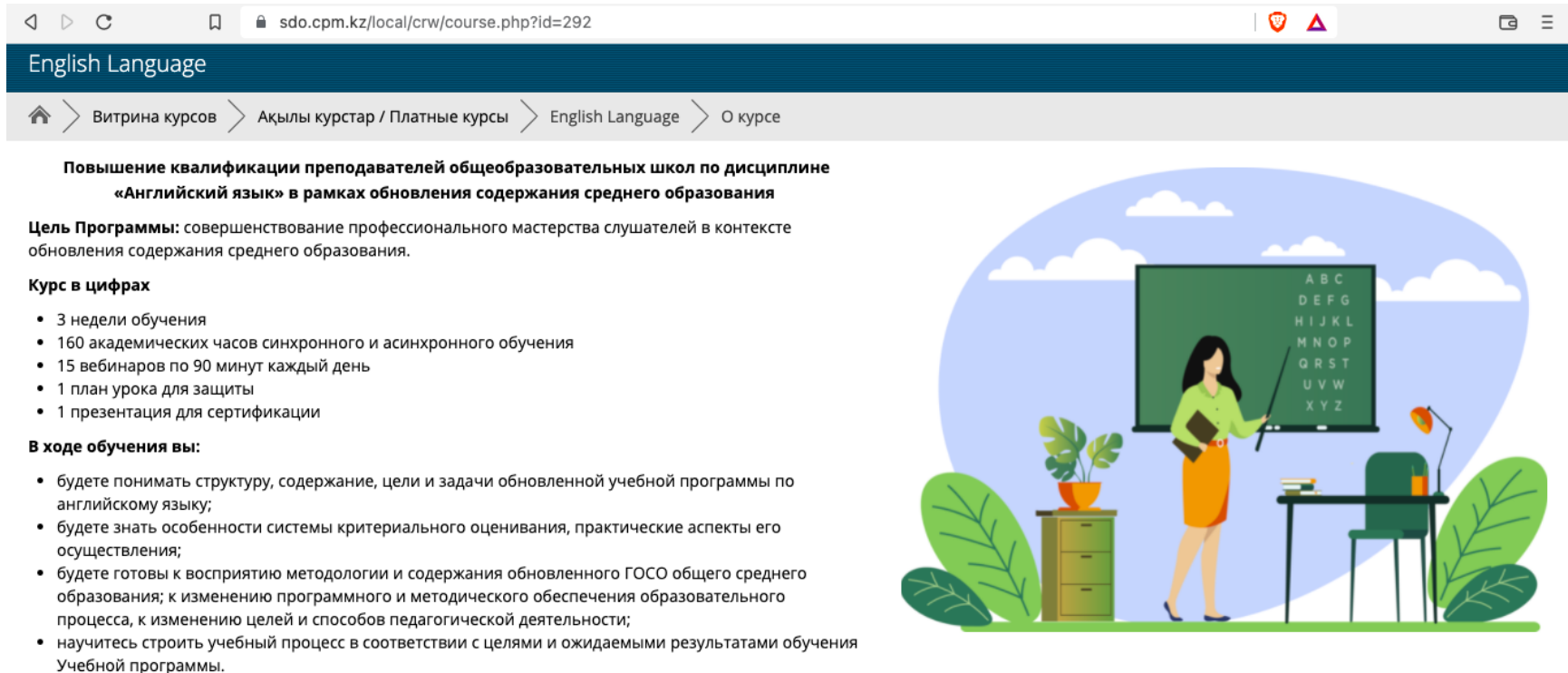
Signature: \_\_\_\_\_ Date: \_\_\_\_\_

The extra copy of this signed and dated consent form is for you to keep.

## Appendix D

Screen captures of Web pages

CPM Course on Updated Curriculum for EFL Teachers (Retrieved from <https://sdo.cpm.kz/local/crw/course.php?id=292>)



English Language

Витрина курсов > Ақылы курстар / Платные курсы > English Language > О курсе

**Повышение квалификации преподавателей общеобразовательных школ по дисциплине «Английский язык» в рамках обновления содержания среднего образования**


**Цель Программы:** совершенствование профессионального мастерства слушателей в контексте обновления содержания среднего образования.

**Курс в цифрах**

- 3 недели обучения
- 160 академических часов синхронного и асинхронного обучения
- 15 вебинаров по 90 минут каждый день
- 1 план урока для защиты
- 1 презентация для сертификации

**В ходе обучения вы:**

- будете понимать структуру, содержание, цели и задачи обновленной учебной программы по английскому языку;
- будете знать особенности системы критериального оценивания, практические аспекты его осуществления;
- будете готовы к восприятию методологии и содержания обновленного ГОСО общего среднего образования; к изменению программного и методического обеспечения образовательного процесса, к изменению целей и способов педагогической деятельности;
- научитесь строить учебный процесс в соответствии с целями и ожидаемыми результатами обучения Учебной программы.



**Translation**

**Title:** The advanced training of teachers at a secondary education school in the discipline "English" as part of updating the content of secondary education.

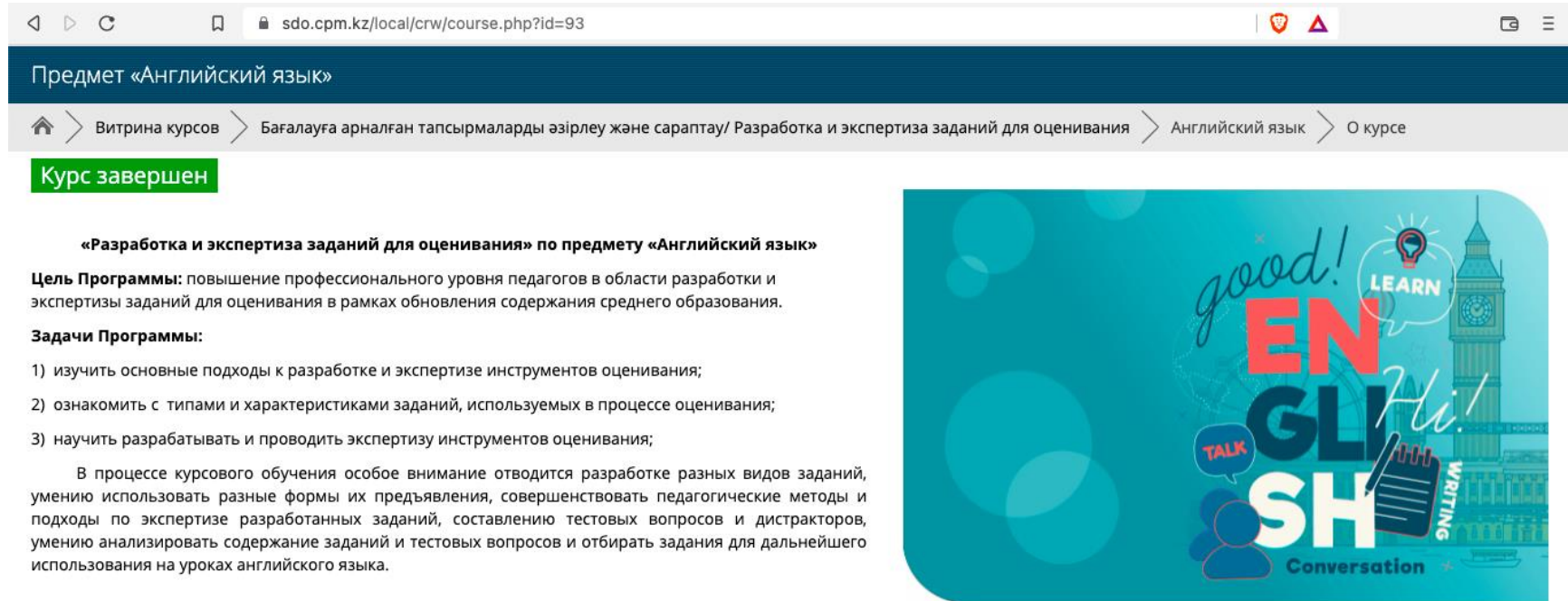
**Program goal:** to improve the professional skills of teachers in the context of updating the content of secondary education.

**The course in numbers**

- 3 weeks of training
- 160 academic hours of synchronous and asynchronous learning
- 15 webinars of 90 minutes each day
- 1 lesson plan for defence
- 1 certification presentation

**During the course you will:**

- understand the structure, content, goals and objectives of the updated English curriculum;
- know the features of the criteria-based assessment system, the practical aspects of its implementation;
- be ready to accept the methodology and content of the updated SES of general secondary education; to change the software and methodological support of the educational process, to change the goals and methods of pedagogical activity;
- learn how to build the educational process in accordance with the goals and expected learning outcomes of the Study Program.

CPM Course on Assessment for EFL Teachers (Retrieved from <https://sdo.cpm.kz/local/crw/course.php?id=93>)

Предмет «Английский язык»

Витрина курсов > Бағалауға арналған тапсырмаларды әзірлеу және сараптау/ Разработка и экспертиза заданий для оценивания > Английский язык > О курсе

**Курс завершен**


**«Разработка и экспертиза заданий для оценивания» по предмету «Английский язык»**

**Цель Программы:** повышение профессионального уровня педагогов в области разработки и экспертизы заданий для оценивания в рамках обновления содержания среднего образования.

**Задачи Программы:**

- 1) изучить основные подходы к разработке и экспертизе инструментов оценивания;
- 2) ознакомить с типами и характеристиками заданий, используемых в процессе оценивания;
- 3) научить разрабатывать и проводить экспертизу инструментов оценивания;

В процессе курсового обучения особое внимание отводится разработке разных видов заданий, умению использовать разные формы их предъявления, совершенствовать педагогические методы и подходы по экспертизе разработанных заданий, составлению тестовых вопросов и дистракторов, умению анализировать содержание заданий и тестовых вопросов и отбирать задания для дальнейшего использования на уроках английского языка.



**Translation**

**Title:** "Development and examination of assignments for assessment" in the subject "English language"

**Program goal:** improving the professional level of teachers in the field of development and examination assignments for assessment as part of updating the content of secondary education.

**Program Objectives:**

1. to study the main approaches to the development and examination of assessment tools;
2. familiarise with the types and characteristics of tasks used in the assessment process;
3. to teach how to develop and conduct an examination of assessment tools;

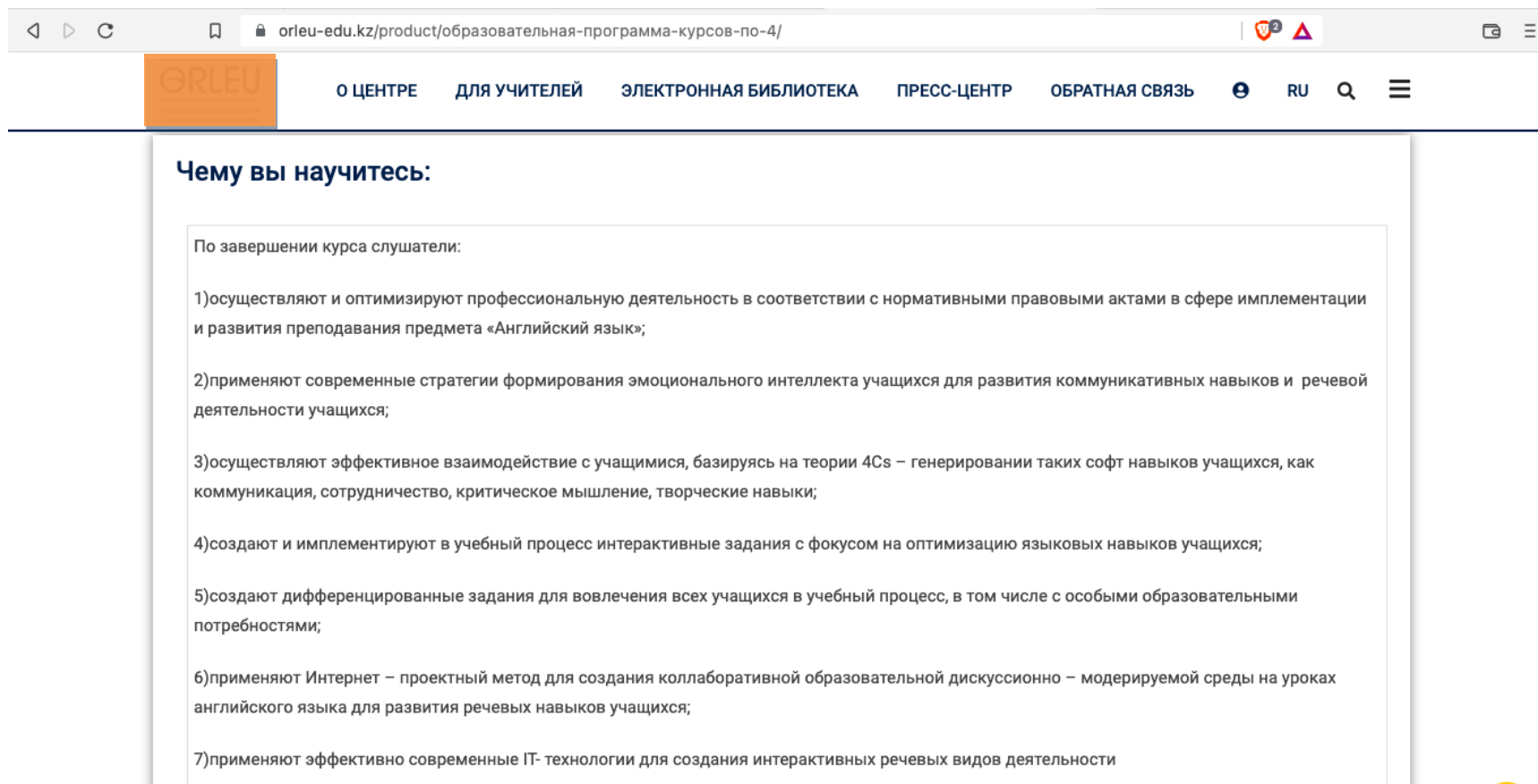
In the process of coursework, special attention is paid to the development of different types of tasks, the ability to use different forms of their presentation, to improve pedagogical methods and approaches for examining the developed tasks, compiling test questions and distractors, the ability to analyse the content of tasks and test questions and select tasks for further use in the classroom. in English.

Orleu PD Course for EFL Teachers (Retrieved from <https://orleu-edu.kz/product/%d0%be%d0%b1%d1%80%d0%b0%d0%b7%d0%be%d0%b2%d0%b0%d1%82%d0%b5%d0%bb%d1%8c%d0%bd%d0%b0%d1%8f-%d0%bf%d1%80%d0%be%d0%b3%d1%80%d0%b0%d0%bc%d0%bc%d0%b0-%d0%ba%d1%83%d1%80%d1%81%d0%be%d0%b2-%d0%bf%d0%be-4/>)

### О курсе:

Форма обучения	Дистанционно в асинхронном режиме, Дистанционно в синхронном режиме, Офлайн (Offline)
Продолжительность	40 часов, 80 часов
Форма завершения	Сертификат
Язык	Русский
Уровни образования	Начальное, основное среднее, общее среднее образование



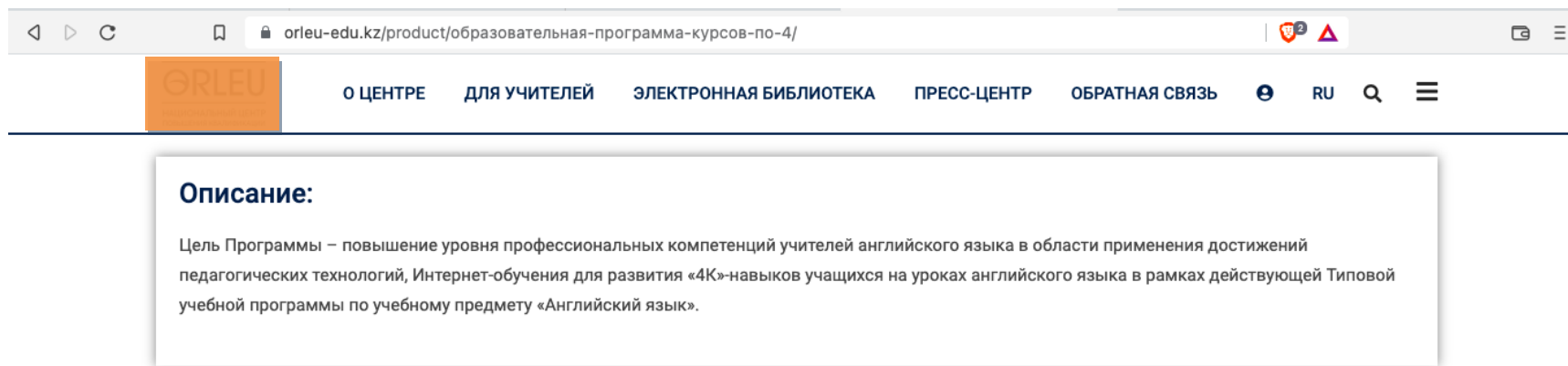


The image shows a browser window with the URL `orleu-edu.kz/product/образовательная-программа-курсов-по-4/`. The navigation menu includes: О ЦЕНТРЕ, ДЛЯ УЧИТЕЛЕЙ, ЭЛЕКТРОННАЯ БИБЛИОТЕКА, ПРЕСС-ЦЕНТР, ОБРАТНАЯ СВЯЗЬ, RU, and a search icon. The main content area is titled "Чему вы научитесь:" and contains a list of seven learning outcomes for course graduates.

**Чему вы научитесь:**

По завершении курса слушатели:

- 1)осуществляют и оптимизируют профессиональную деятельность в соответствии с нормативными правовыми актами в сфере имплементации и развития преподавания предмета «Английский язык»;
- 2)применяют современные стратегии формирования эмоционального интеллекта учащихся для развития коммуникативных навыков и речевой деятельности учащихся;
- 3)осуществляют эффективное взаимодействие с учащимися, базируясь на теории 4Cs – генерировании таких софт навыков учащихся, как коммуникация, сотрудничество, критическое мышление, творческие навыки;
- 4)создают и имплементируют в учебный процесс интерактивные задания с фокусом на оптимизацию языковых навыков учащихся;
- 5)создают дифференцированные задания для вовлечения всех учащихся в учебный процесс, в том числе с особыми образовательными потребностями;
- 6)применяют Интернет – проектный метод для создания коллаборативной образовательной дискуссионно – модерлируемой среды на уроках английского языка для развития речевых навыков учащихся;
- 7)применяют эффективно современные IT- технологии для создания интерактивных речевых видов деятельности



orleu-edu.kz/product/образовательная-программа-курсов-по-4/

ORLEU

О ЦЕНТРЕ    ДЛЯ УЧИТЕЛЕЙ    ЭЛЕКТРОННАЯ БИБЛИОТЕКА    ПРЕСС-ЦЕНТР    ОБРАТНАЯ СВЯЗЬ    RU    🔍

**Описание:**

Цель Программы – повышение уровня профессиональных компетенций учителей английского языка в области применения достижений педагогических технологий, Интернет-обучения для развития «4К»-навыков учащихся на уроках английского языка в рамках действующей Типовой учебной программы по учебному предмету «Английский язык».

**Translation**

**Course Title:** Development of professional competencies of an English teacher

**About the course:**

Form of study	Remote in asynchronous mode, Remote in synchronous mode, Offline
Duration	40 hours, 80 hours
Completion Form	Certificate
Language	Russian
Levels of Education	Primary, basic secondary, general secondary education

**What will you learn:****Upon completion of the course, teachers will:**

1. carry out and optimise professional activities in accordance with regulatory legal acts in the field of implementation and development of teaching the subject "English";
2. apply modern strategies for the formation of students' emotional intelligence to develop communication skills and speech activity of students;
3. carry out effective interaction with students, based on the theory of 4Cs - the generation of such soft skills of students as communication, cooperation, critical thinking, creative skills;
4. create and implement interactive tasks in the educational process with a focus on optimising the language skills of students;
5. create differentiated tasks to involve all students in the educational process, including those with special educational needs;

6. use the Internet - the project method to create a collaborative educational discussion - moderated environment in English lessons to develop students' speech skills;
7. effectively apply modern IT technologies to create interactive speech activities

**Description:**

The purpose of the Program is to increase the level of professional competencies of English language teachers in the field of applying the achievements of pedagogical technologies, Internet learning for the development of "4K" skills of students in English lessons within the framework of the current Standard Curriculum for the subject "English Language".

## Appendix E

### Interview Transcripts

#### *Interview 1 In-Service Rural School Teacher (Transcribed from Kazakh)*

Categories	Interview Questions	Answers
<b>Warm-up Questions</b>	<p>1. How long have you been teaching English at school?</p> <p>2. Were you involved in state-funded teacher training from organisations like Orleu or Center of Excellence?</p> <p><b>Probes:</b> What types of interventions did you participate in? (Training, seminar, workshop etc.)</p>	<p>I started working at school in 2014.</p> <p>Yes, in 2016, in June. It was an entire course by CoE. After that time, I took training on critical thinking and the level three on the updated curriculum by Orleu. The trainer was a teacher of the third category from our school.</p>
<b>What are teachers' attitudes and beliefs about language assessment in classrooms?</b>	<p>3. Did you work at school before the updated curriculum reform was introduced?</p> <p>What assessment models did you use at that time?</p> <p>Do you see any differences or similarities between the traditional and updated assessment models?</p>	<p>We used a traditional assessment method in 2014, day-to-day monitoring and setting a final mark depending on overall progress.</p> <p>In the beginning, I, as any young teacher, had my mentor. She helped me deal with daily tasks like preparing lesson plans and planning assessments, mostly reading and writing. Assessment of speaking and listening skills was rare, and we did listening tasks about twice or once a week. And the reading and writing tasks were utilized in almost every lesson. In short, we lacked a focus on listening and reading tasks.</p> <p>I cannot say there are no similarities because we assess the same skills: reading, writing, listening and speaking. But, I think the updated model is more effective because it requires assessing all four skills where you cannot skip any</p>

of them, using points instead of marks, where formative and summative tests accumulate the final results. I find it convenient and time-saving in comparison to the old procedures.

**Did you have other difficulties in traditional assessment?**

As I already mentioned, the traditional assessment did not require assessing all four skills, so listening and speaking were mainly totally skipped, but it is not possible in the updated assessment model. In addition, in updated form, students get points to compete to achieve higher results, not only for final marks, and they see final marks only at the end of the term.

**What grades do you teach?**

I teach predominantly primary schools, from 1 to 4, and some grades from secondary.

In addition, it is easier to put a mark connected to students' level when you have points instead of marks. For example, in traditional, we had only marks 2, 3, 4 and 5, where we also mostly avoided using mark 2, but in a formative assessment, we have more points, which is more effective and more manageable in marking.

4. What do you think about the teachers' assessment of literacy?

How does it affect their professionalism?

How does it affect the language teachers' professionalism?

5. What is your self-perceived level of language assessment literacy (LAL)? What do you think are the reasons for your answer?

**Probes:** What is the most difficult aspect of

There are many types of teachers with different thinking and abilities. I can't decide whether someone is good or bad at assessing. I mean, any teacher needs experience, and self-development is achievable to any of us. There is no limit to learn something new.

Regarding my own state, I can say that I am somewhere in the middle because we are teachers of an ordinary rural school. In addition, we do not have students who know advanced English, as in some urban schools. I always try to advance my English teaching skills because we are the

---

LAL in your classroom practice? Why is this?

ones who learn and improve them all the time.

**What are the other factors that affect your self-confidence in assessing?**

The engagement of students. Most of our students are not interested in learning English as they find it too difficult.

We, as teachers, understand it and try to improve the situation; however, sometimes, it does not work.

We have some good students who participate and win in Olympiads, but their number is low.

Another factor is the lack of attention to teachers from rural schools. I have worked here for three years and did not participate in any seminar or workshop. Moreover, teachers need to observe lessons in other schools to exchange ideas and practices; however, it is not happening here. I think it affects our professionalism. My previous school was urban; we were used to attending English clubs, which was very useful for me.

**Could you tell me more about English clubs, what place is it?**

English club is the intervention where different teachers of English gather at, for example, a library to ask questions on different topics, mostly about their classroom practices.

We do not have such clubs here, and it could help us develop our teaching and language skills. Moreover, I believe that improving speaking skills among teachers of rural schools could improve teaching English. So, preparing seminars on this problem should be the main focus of schools administrators for us. We do not have the time and power to organize such events ourselves.

**Do you have challenges in assessing?**

Of course, I have. I teach all four grades of primary school, and teaching them to read is difficult for me. I think it is not suitable for my work because teaching and assessing

---

them takes much time and energy. It could be much easier if I had some grades from primary school and some from secondary, for example. Our vice principal does not understand such situations, as she does not teach English to kids. If administrators of schools understand the amount of workload we have, it would be better. Imagine that only one teacher is responsible for teaching all primary classes, more than 100 students overall. So, the quality of teaching suffers.

**To what extent you are confident when you set a particular mark on a student? Do you have any instructions that help you?**

There are no detailed handbooks for us. As a teacher who has worked with the same children for months, I know their approximate levels. Moreover, I take most of the exercises from a book specially designed for a specific grade, e.g. Excel for grade 5, Smiles for primary school, so I can be for at least 60% sure that my assessment is fair.

**What are the teachers' suggestions about language assessment in classrooms?**

6. What do you think should be done to improve the teachers' LAL in the school you work in? What about at the national level?

**Probes:** What kind of PD interventions would be useful to you?

I think teachers' collaborative work at school is essential and can help in many situations. For example, they can organise English clubs, and there are also teachers of other subjects who learn or speak English to join it. At a school level, it can be beneficial to improve English in general, which will improve assessment too.

Regular seminars or contests on a needed theme are usually helpful at the national level. However, I can add that there are international teachers at any city in KZ, so the Ministry of Education can support ideas in organising events where those native speakers come to rural schools with courses, seminars or other exciting things.

Moreover, I think the assessment itself is not only about students' abilities in using language, but it is also about

---

their discipline, preparedness, focusing, e.g. doing homework, etc. In addition, finally, it can be challenging for a teacher to assess student's performance when the teachers themselves have poor English; it is a fact.

---

*Interview 2 In-Service Urban School Teacher (Transcribed from Russian)*

Categories	Interview Questions	Answers
<b>Warm-up Questions</b>	<ol style="list-style-type: none"> <li>1. How long have you been teaching English at school?</li> <li>2. Were you involved in state-funded teacher training from organisations like Orleu or Center of Excellence?</li> </ol> <p><b>Probes:</b> What types of interventions did you participate in? (Training, seminar, workshop etc.)</p>	<p>From 2015, about six years.</p> <p>CoE regularly provides various master classes both online and offline. Therefore, the abbreviation of CoE is constantly circulated in our talks, they always work with our school, and we often take part in their events.</p> <p><b>Tell more about the events' format and topics, please.</b></p> <p>The quick thing that I remember recently is an online seminar on soft skills. I remember it because it was different from what they usually do, different from traditional/professional teaching seminars. In general, we are provided with some short online or face-to-face lectures on teaching tips, like ways to improve remembering the materials of lessons, working with presentations etc.</p>
<b>What are teachers' attitudes and beliefs about language assessment in classrooms?</b>	<ol style="list-style-type: none"> <li>3. Did you work at school before the updated curriculum reform was introduced?           <ul style="list-style-type: none"> <li>What assessment models did you use at that time?</li> <li>Do you see any differences or similarities between the traditional and updated assessment models?</li> </ul> </li> <li>4. What do you think about the teachers' assessment of literacy?</li> </ol>	<p>In 2014, the teaching program at the school where I worked already updated, and in 2015 only the higher grades had a formative assessment in older variant, which was with ratings like "achieved, not achieved, achieving". Later, all the grades at our school had a similar updated program which we have today at all schools in Kazakhstan. So, I did not work at school with the old traditional assessment system.</p> <p>Our school pays special attention to assessment. It is an essential aspect. However, although we have various seminars, not often on such themes particularly, AND some digital materials shared, it is often not learnt independently; OR when there are some questions on it, it often fails to</p>

---

<p>How does it affect their professionalism? How does it affect the language teachers' professionalism?</p>	<p>ask... Therefore, I think many teachers want to know more about the assessment system because most of their knowledge on assessment is a bit superficial. Namely, they know the basics and some practices; however, if it is about explaining it to someone professionally, I think most of them will face problems.</p>
<p>5. What is your self-perceived level of language assessment literacy (LAL)? What do you think are the reasons for your answer?</p>	<p>If a teacher is professionally informed on the assessment system, it will generally help them check summatives and formatives or other works, explain their marks to students and other colleagues, why and how they have a specific rating, and improve the teaching. So teachers and students could work in tandem when the process is more evident to both, and its impact is only positive.</p> <p>I think it is a bit lower than medium. Even though I know a lot about criteria-based assessment and some other basics, mainly due to lack of time, the process of assessment happens in the form of a quick review. It can be explained that the process itself is labour-intensive and time-consuming. Especially when we mark essays, we face various specifics of language teaching related to the age of students and features of the criteria-based system that may affect the complexity and the quality of assessment. Therefore, my level of LAL is lower than medium or close to the middle.</p>
<p><b>Probes:</b> What is the most difficult aspect of LAL in your classroom practice? Why is this?</p>	<p>If we take some technical moments in assessing all four language skills of students, we should use a specific approach to each of them. For example, IELTS system, ... if it is about writing or reading, works of students in higher grades are assessed differently. As a result, you have to master much information and skills to conduct a proper assessment and provide good feedback. Therefore I think it is especially problematic because we usually focus on assessing grammar or lexis; however, other aspects like content are</p>

---

assessed by guesswork. I mean, we rely on the rich experience that allows us to look at students' work and decide what level or mark is appropriate in different situations, but if there is a need to explain our decision professionally with good feedback to a student or school administration, I, personally, may have some challenges. In other words, it isn't easy to do detailed reports, as we must pay attention to many factors and aspects, and there are different specifics in different grades, namely, primary, secondary or higher, they all have specific things, we should learn them, and ... I think we need some pieces of training not less than once a year, namely, to share an experience or to attend seminars. In other words, we need a second opinion.

**What are the teachers' suggestions about language assessment in classrooms?**

6. What do you think should be done to improve the teachers' LAL in the school you work in? What about at the national level?

**Probes:** What kind of PD interventions would be useful to you?

I think sharing ideas is the most optimal variant; listening to your colleagues who have already mastered it to a certain level. And the views of teachers from other places is also essential, so we should combine them all: the knowledge from external sources and local practices. The important thing is to find an appropriate time and not to bring more work to teachers, for example, additional reports. And it should have a more accessible and concise format, brief, only containing the primary information. As a result, every teacher may have their own notes or a handbook, and maybe some practical exercises that are not time-consuming. In addition, we annually have different contests at a national level for teachers. So, we have winners, and maybe they could produce valuable materials, not just academic articles, which can be boring for the majority, but maybe some exciting presentations, of course, unobtrusive, lite, with elements of humour, etc., focusing on masses. It could be beneficial.

*Interview 3 Pre-Service Teacher (Transcribed from Russian)*

Categories	Interview Questions	Answers
<b>Warm-up Questions</b>	<p>1. How long have you been teaching English at school?</p> <p>2. Were you involved in state-funded teacher training from organisations like Orleu or Center of Excellence?  <b>Probes:</b> What types of interventions did you participate in? (Training, seminar, workshop etc.)</p>	<p>I did not teach at school; I only taught at language centres for some time.</p> <p><b>What is your degree?</b>  I finished university with a diploma in two foreign languages, English and Chinese.</p> <p>No, it didn't happen for some reason. We were studying in the program that required last year students to pass an examination on the qualification of teachers, and that practice was just being implemented. But the Covid-19 lockdown influenced the process, and everything was cancelled. Initially, we were supposed to get free courses in English and tests, but unfortunately, we did not get any preparations or online tests. I graduated in 2020.</p>
<b>What are teachers' attitudes and beliefs about language assessment in classrooms?</b>	<p>3. Did you work at school before the updated curriculum reform was introduced?  What assessment models did you use at that time?  Do you see any differences or similarities between the traditional and updated assessment models?</p> <p>4. What do you think about the teachers' assessment of literacy?</p>	<p><b>What assessment models do you know?</b>  When I was at school, we had 5-scale marks. However, at university, we had a course on the updated model of assessment (I do not remember the exact name of the course), where we learned the theory and principles of the new assessment model. Further in year 3 and at the beginning of the last year, we attended schools for short internships where we had lessons, prepared and conducted some small tasks and tests, saw the main processes of formatives and summatives. However, I did not get much information and knowledge on assessment practises.</p> <p>I believe that it is a human factor. It is perhaps dependent on teachers' knowledge and their human qualities.</p>

---

<p>How does it affect their professionalism? How does it affect the language teachers' professionalism?</p>	<p>Sometimes teachers should be sufficiently competent to raise or lower their level of requirements to achieve effective teaching. Marks of students are also indicators of teachers' proficiency.</p> <p>It has a significant impact on a teacher's work at school because parents need to know the level of their children, the school needs to see the teacher's work, and the students want to know about their progress, so maybe this is the reason for schools to have such rating system to assess everything. Teachers need to know the principles and practices of assessment for 90% because it is not possible to do it for 100%. All the knowledge, even the theories from books, are not always practical, there are various situations in classrooms, and it is hard to predict them and act as it is instructed in documents.</p>
<p>5. What is your self-perceived level of language assessment literacy (LAL)? What do you think are the reasons for your answer?</p>	<p>Maybe, it is about the normal or under the normal, and I don't have any experience of working at school. Most likely that in the beginning, I will have some challenges with the systems of assessment etc., but after a while, I hope I will be able to have some progress.</p> <p>If I am provided with teaching and assessment materials with tasks and instructions especially prepared for specific school age, I think I would be 80-90% sure that I teach and assess in the right way.</p>
<p><b>Probes:</b> What is the most difficult aspect of LAL in your classroom practice? Why is this?</p>	<p>I didn't work at school, but when I imagined it, I thought it would be difficult for me to assess speaking. Because students are of different language proficiency levels, also their personalities are different too. So sometimes, it can be hard to make them speak and answer the questions even when they know the answers.</p>

---

---

**What are the teachers' suggestions about language assessment in classrooms?**

6. What do you think should be done to improve the teachers' LAL in the school you work in? What about at the national level?  
**Probes:** What kind of PD interventions would be useful to you?

I believe both teachers and students should be provided with accurate information on the processes; teachers should inform students about skills and knowledge to be acquired by the end, of course, etc., while teachers should be instructed on the steps and actions for some unexpected situations; well, e.g. to make teachers know that there are different students with different needs, etc. Also, they should use different methods. I can be mistaken because I didn't teach at school in front of big classes with over 16 students in a classroom. Normally, when you teach big groups, one or two students have challenges in learning the language in the beginning, maybe even with some physical disabilities, so teachers should be cautious about that. In addition, professors in universities provide students with a syllabus and a list of requirements where everything is described, so I think a similar practice can be beneficial at schools. It can also be beneficial in preparing school students for the university lifestyle.

At the national level, I believe that the ministry of education should work on delivering clear and detailed instructions because we see that teachers are not always informed about the requirements that governments set for their teaching. For example, we had different reforms and changes in the educational system, but they were not correctly presented, so teachers misinterpreted them, and finally, students faced problems.

For teachers like me, pre-service EFL teachers, there are various training courses, webinars and other interventions for professional development. However, I would acknowledge the need for some practical events or courses where future teachers can see the teaching process and participate in it. Additionally, well-prepared

---

---

mentors at schools who are informed and experienced in most schoolwork could be helpful. They could advise or guide new teachers, and it would help to learn fast and effectively.

---

