

UDC 378.147:811.111'243]:371.263

Olga Kvasova (Ukraine)

ORCID ID: 0000-0002-1479-0811

Yuliia Kaleniuk (Ukraine)

ORCID ID: 0009-0001-7506-4194

A PREVENTIVE INTERVENTION IN THE LANGUAGE TESTING AND ASSESSMENT COURSE: BUILDING PRE-SERVICE TEACHERS' MULTIPLE-CHOICE ITEM WRITING SKILLS

The article addresses the development of pre-service teachers' language assessment literacy within the course "Testing and assessment of Learning Outcomes", which is offered at the bachelor's level in Taras Shevchenko National University of Kyiv. Constructing test tasks for classroom-based assessment poses certain challenges to students, especially with respect to the format "Multiple Choice Questions" (MCQs) for the assessment of careful reading skills. This format is by far the most well known in language education, however, writing quality MCQs is time and effort consuming, and administration of low-quality teacher-constructed MCQs may yield unreliable results of learning the language. To mitigate conceptual, linguistic and technical difficulties, which have been observed during the years of teaching the course, a preventive intervention was collaboratively designed by the course instructor (the mentor) and a research-minded master's level student (mentee), who further conducted the intervention in the classroom. The intervention included preparatory activities, tasks conducted in the classroom, home assignment and a performance test. The results of the performance test were complemented by the data obtained via a feedback questionnaire, which elicited students' perceptions of the intervention and its impact on their learning. The collation and interpretation of both sets of data demonstrated overall effectiveness of the intervention and its positive evaluation by students.

Keywords: *preventive intervention, language assessment literacy, pre-service teachers, multiple choice questions task design.*

Background

The overall goal of teaching consists in enabling good learning outcomes although it also aims at making learning process engaging, thought-provoking and stimulating learner autonomy and responsibility. To this end, the majority of teachers tend to utilize strategies that can maximize learning, innovate teaching methods and techniques, employ digital technologies and AI to make learning up to date. Reflective teachers usually think over their classroom experiences, especially if they observe their students struggling with some aspects of learning. Identification of misunderstanding, misuse of material and inability to improve may signal that teachers should introduce changes in their instructional practice and remedy learning. In this line of thought, any teacher's action that modifies current practice to eliminate learners' problems and improve learning is termed as intervention.

The concept of intervention is not new although the 2020s have witnessed a growing amount of research into its use in education – in particular subjects, e.g., foreign languages, and in overall teacher education. This article addresses the use of intervention to mitigate students' challenges in item-writing within the course "Testing and assessment of learning outcomes" (TALO). The course has been delivered for students majoring in Ukrainian language and literature an English for over a decade and has undergone several modifications. However, instructional practice testified that students continued to struggle with some tasks set before them. In the article, we describe how the difficulties were identified, and what actions were planned and implemented to mitigate these difficulties. The study employed both quantitative and qualitative research tools to obtain the necessary data and interpret them.

To ensure that interventions are highlighted well enough "for the reader who wants to gain insight in its intricacies and thereby find out what its crucial elements could have been" (Rijlaarsdam et al., 2018, p. 280 cited in Koster & Bouwer, 2018) we described the intervention in full detail. As Koster & Bouwer (2018) pointed out, a detailed description of an intervention is also important from a practical point of view, to foster dissemination and successful implementation of the intervention into practice. The article contains a preparatory exercise, two tasks that integrate teacher explanation of rules of constructing quality items with students' hands-on exploration of good and faulty items, and a home assignment. It concludes with the presentation and interpretation of students' scores on exit performance tests, and analysis of responses to a feedback questionnaire.

Literature review. Interventions were initially used in psychological studies addressing problematic behaviour (academic, socio-emotional or other) of learners, aiming for improvement or change. Today, due to educational reforms going on in various contexts, large scale intervention programmes are carried out to address innovations in instruction and their effects on teaching and learning. Intervention studies are typically prompted by a mismatch between innovative curricular requirements and conventional teaching process, therefore, to be implemented in practice, the innovation must be positively evaluated with respect to its effectiveness. In this case, interventions are rigorously structured and performed under scrupulous control (König et al., 2025).

Another type of intervention is located on the level of teaching a particular subject to a particular cohort of learners, who struggle with particular content point of a curriculum. Classroom interventions are subdivided into corrective ones, which aim to remedy existing problems of learners, and preventive, which aim to mitigate difficulties

that the learners are likely to encounter while developing new skills. Preventive interventions are reported to develop higher order thinking skills such as inferencing, creativity, and critical thinking and therefore are more impactful.

Both types of interventions aim at improving instructional quality and supporting students' learning progress, with this goal being dependent on approaches and techniques employed by practitioners. This statement is in line with existence of an effect chain determined by researchers of teacher education (e.g., Desimone, 2009). This chain is activated by an increase in teacher competence ensuing training or self-study and conscious goal setting, after which the changed practice produces positive effect on the quality of teaching and the learning achievements of students. Despite calling this assumption plausible, König et al. (2025) question the actual effectiveness of this chain claiming that scientific knowledge of learning effectiveness is still limited and needs further investigation.

König et al. (2025) also differentiate concepts of intervention applied to formal teacher education and teaching-level interventions defining the latter as "approaches that apply evidence-based teaching concepts or practices to be (learned and) applied by teachers, possibly in combination with evaluating implementation for teaching concept or practice (and their students' learning)". This definition suggests that teaching-level interventions should be based on evidence of students' grasp of some teaching concept, apply informed strategies to teach this concept and evaluate the results of students' acquisition of the concept taught.

Stranovská & Ficzer (2022) assert that interventions should be properly designed and implemented. In this case, "interventions can be a valuable tool for teachers allowing them to utilize various teaching methods, learning strategies, and teaching styles (such as active, social,

project-based learning, direct and indirect methods) or to replace traditional techniques with more innovative ones" (Stranovská et al., 2025, p.3867). To design an intervention properly, Kalinowski et al. (2019) offer a framework that includes *content, structure, and didactic*: the content of the intervention (topical focus and its area), the duration as a structural feature, and basic characteristics indicating didactic features (provision of coursework, transfer into teaching practice, using digital resources). Clearly, such design looks quite comprehensive. We would only add *evaluation* of learning outcomes which will not necessarily follow the classical 'pre-post' design but use some other measure.

Intervention as a strategy affecting learners' achievements in a particular topical area seems to be high in demand. The past decade has seen publications about interventions in teacher education (Barrio et al., 2015; Adams et al., 2024; Wohlfart, 2024; Alfonso et al., 2025) and in L2 learning (Koster & Bouwer, 2018; Sato & Lara, 2019; Kalaitzi & Theocharis, 2023; Stranovská et al., 2025), the areas being closely related to the topic of our investigation. The listed studies differ in structure/duration and in didactic features providing versatile insights for researchers of interventions. However, we did not find any research studies focusing on interventions within particular professional courses for prospective L2 teachers, in particular courses in language testing and assessment (LTA).

The issue of developing pre-service teacher language assessment literacy (LAL) has received its focus quite recently in the studies by Volante & Fazio (2007), Fröjdendahl (2018), Ukrayinska (2018), Giraldo & Murcia, (2019), Kvasova (2022a). This literacy develops within initial teacher education when prospective teachers have little or no experience of real-life instructional practice. Thus, LAL of prospective teachers should not be confused with the LAL of practicing teachers

(Malone, 2017). Moreover, it is fully context-dependent (Scarino, 2013; Tsagari, 2017), and can be specified only regarding its relevance to, and consistence with the curriculum of the course taught, which in turn is aligned with learner's background and maturity, as well as the purpose of their studying LTA (Kvasova, 2020).

Despite certain differences in the construct of practicing and pre-service teachers' LAL, there are some areas in LTA which are similarly challenging for both cohorts of teachers. We refer to item writing which is key to the development of valid test tasks (Rossi et al., 2025) both in standardized and classroom-based assessment and, therefore, should be properly explained to trainees and followed by intensive practice. Item writing occupies a central position within teacher training courses in LTA (Tsagari et al., 2018; Ukrayinska, 2018; Kvasova, 2022b). However, both linguistic (Rossi, 2021; Ukrayinska, 2025) and technical quality of items constructed by pre-service teachers require additional effort from course instructors. This conclusion emphasizes the necessity to implement a preventive targeted intervention within the course TALO.

The Research Questions posed in this study are as follows:

RQ1. Was the intervention aimed at preparing students for constructing item 'Multiple choice questions' effective?

RQ2. What was the students' evaluation of the intervention and its impact on item writing?

The study was conducted at Taras Shevchenko National University of Kyiv, Ukraine.

Methodology

Procedure

The current study was carried out collaboratively by two researchers – a mentor (a course instructor) and a mentee (a master's student). The preventive intervention, which targeted skills of writing MCQs, was designed by the mentee under the mentor's guidance upon which the mentee (henceforth termed "teacher")

implemented the intervention in the TALO course classroom. The research was performed through the following stages:

1. identification of potential difficulties the students may face.
2. developing series of exercises aimed at mitigating the identified potential difficulties;
3. conducting the intervention in the classroom;
4. testing students' performance in writing items;
5. administering a questionnaire to elicit feedback from students.

The intervention was conducted within one month in Term 1 in accordance with the academic plan of the discipline TALO.

Participants

The data were collected from eight students in the 3d year group of students trained to be teachers of Ukrainian language and literature, and English. They all were female, aged 19-20, native speakers of Ukrainian.

Methods

The current study adopted a mixed-methods approach. A performance test to obtain numerical scores as evidence of effectiveness of the intervention was designed and administered by the teacher. It aimed to evaluate five MCQs constructed by students against a 10-point holistic scale.

An online questionnaire was administered to capture students' perceptions/evaluation of the intervention's effectiveness. The questionnaire aimed to capture students' evaluation of intervention usefulness (preparatory task and tasks 2 and 3), the encountered challenges, and the clarity of the instructional materials. To ensure authentic and detailed responses, the questionnaire was administered in the students' native language (Ukrainian). Both quantitative and qualitative data were collected through a combination of Likert-scale items and open-ended questions. The questionnaire consisted of two parts: a checklist of ten "I have learned..." statements aligned with the key item-writing

principles adapted from Coombe et al. (2007), rated on a 5-point Likert scale (1 = "strongly disagree", 5 = "strongly agree"), and five open-ended questions inviting students to comment on specific aspects of the workshop and their future application of the acquired skills. The questionnaire was administered online via Google Forms.

Findings and discussion

In this section, we will first describe intervention stages 1-5 mentioned above focusing on the activities included in the intervention, then provide the results of students' performance test and in the end discuss the data obtained through the questionnaire.

Intervention "Training students in developing good MCQ items"

The topic chosen by the teacher for the intervention – writing multiple choice questions (MCQs) to test reading – is considered demanding not only for pre-service teachers but for experienced teachers as well. The in-depth analysis of MCQs constructed by students in previous years revealed the typical mistakes students tend to make. These include writing options of different lengths, writing unparallel options, writing options that lift verbatim chunks of original text, writing implausible distractors and writing distractors which are off the context. These violations are well described by Coombe et al. (2007). In the intervention design, the strategy to compare and contrast good and faulty items appeared constructive: a required number of juxtaposed items was selected from an anonymous bank of students' works and further presented in the tasks as pairs to compare and contrast. This juxtaposition enabled this year's cohort to deduce the rules for writing good items by themselves, which resulted in better interiorization of the rules and their accurate use in constructing own items.

The intervention involves a preparatory task, tasks 2 and 3 conducted collaboratively in class, a home assignment and a performance test.

Preparatory task 1A. Read a short text about an embarrassing moment in Niamh's life. Then choose correct answers A, B, C, or D for questions (1-5) in the box. Note if the items are clear / confusing / complicated.

Niamh, Devon: "I was in school and as usual I was swinging back on my chair,

really annoying the teacher. However, I ended up losing my balance and I fell straight out the fire exit door behind me! Everyone in the class was in hysterics and when I got up, the teacher gave me a huge row! Crazy cringe or what!"

1. Niamh

- A** had a great fall
- B** got stuck in the fire exit
- C** was in hysterics
- D** broke his chair

2. Niamh fell out the fire exit door because

- A** he was trying to escape
- B** he was annoying the teacher
- C** he was playing a game
- D** he lost her balance

3. What happened after falling?

- A** The teacher applauded the narrator's balancing act.
- B** The entire class sympathized with the narrator.
- C** The narrator fell out the fire exit door, causing laughter in the class.
- D** The teacher ignored the incident and continued with the lesson.

4. While swinging in his chair, Niamh... .

- A** annoyed the teacher
- B** lost his balance
- C** became hysterical
- D** fell straight forward

5. How did the classmates react to the incident?

- A** They laughed hysterically
- B** They were concerned and helped the speaker
- C** They ignored it completely
- D**) It is not mentioned in the text

Preparatory task 1B. Go over the items again. Decide: Which of the items ... ?

- is/are correct
- do not have a correct answer
- are grammatically inconsistent
- contain options that begin with the same word

contain verbatim chunks from the source text

- is/are written in language more difficult than the text itself
- contain options of different lengths
- stand out graphically

Task 2. Follow the teacher presenting the rules for writing good MCQs. Compare and contrast examples of MCQs in (a-i). Identify violations of the rules in each case and suggest improvements to the items.

(a) A good MCQ for testing reading contains four options.

<p>Suzi felt embarrassed because ...</p> <p>A she misread the price B she forgot her purse C the jeans didn't fit</p>	<p>Suzi didn't buy the jeans because ...</p> <p>A they were too expensive B she forgot her money C she misread the price D the shop assistant was rude</p>
---	--

(b) One option in a good MCQ is correct, two distractors are plausible – one is too general, and the other is too narrow. The incorrect option may be taken off context.

<p>What happened when Niamh leaned back on her chair?</p> <p>A He dropped her book on the floor. B He hit another student. C He broke the chair. D He accidentally fell out of the classroom</p>	<p>Niamh fell out the fire exit door because ...</p> <p>A he was trying to escape B he was careless C he broke the chair D he lost his balance</p>
--	--

(c) The options in a good MCQ are of similar length.

<p>How did the classmates react to the incident?</p> <p>A They laughed hysterically B They were concerned and helped the speaker C They ignored it completely D It is not mentioned in the text</p>	<p>Niamh fell because ...</p> <p>A he leaned back too far B the chair broke C he slipped on the floor D someone pushed him</p>
---	--

(d) The options in good MCQs are written in clear language, of similar difficulty with the text itself, and appropriate for the level of learners.

<p>Suzi felt uncomfortable because ...</p> <p>A the jeans were much more expensive than she thought B she had to leave the store without the jeans due to a mistake C the jeans were out of stock and she couldn't buy them D the price mistake caused her to make an awkward excuse</p>	<p>Suzi was embarrassed because she ...</p> <p>A couldn't afford the jeans B misunderstood the price tag C forgot her money at home D liked the expensive jeans</p>
--	---

(e) The options in good MCQs do not give away the correct answer.

<p>Where was Suzi when she discovered the inexpensive pair of jeans?</p> <p>A At home. B At a party. C Out shopping. D At work.</p>	<p>Suzi felt embarrassed because ...</p> <p>A she misread the price B she forgot her purse C the jeans didn't fit D she spent too much</p>
---	--

(f) In a good MCQ, a word that is repeated at the beginning of all options is placed at the end of the stem.

<p>The girl was ashamed because ... A she had taken a book about sexual diseases B she was late for class C she left the book at home D she forgot the teacher's name</p>	<p>Orlando Bloom fan was embarrassed because she A had to read a book she didn't like B forgot to bring a book to class C brought an inappropriate book by mistake D had no time to choose a book</p>
--	---

(g) In a good MCQ, all options are parallel and grammatically correct.

<p>The problem was that the girl A lost her purse B read the price wrongly C shop assistant rung up another jeans D had already been out of money</p>	<p>Suzi went to the store to buy a new pair of jeans and she A spent £ 100 on new jeans B couldn't find her purse C bought a lovely pair of jeans D couldn't buy a trendy pair</p>
--	---

(h) In a good MCQ, the options do not repeat words/phrase verbatim from the text.

<p>Niamh fell out the fire exit door because A he was trying to escape B he was annoying the teacher C he was playing a game D he was swinging on her chair</p>	<p>Niamh A annoyed teacher B fell from his chair C quarreled with his teacher D got a bad mark</p>
---	---

(i) In a good MCQ, there is one correct answer.

<p>The teacher was angry because Niamh A wasn't listening to him B tried to skip the lesson through the fire exit C had been laughing with the class D had been irritating her</p>	<p>The teacher A caught Niamh playing on his phone B didn't allow him to swing on his chair C laughed hysterically after his fall D penalised him for misbehaviour</p>
---	---

Task 3. In teams, write your MCQs to check understanding of the following text. Then ask your classmates to pilot the MCQs and give feedback for you to improve them.

THE ROCK GARDEN OF CHANDIGARH

In Chandigarh, a city in northeast India, there is an amazing garden. There are tall rocks and small paths between them. There are waterfalls, pools, and hundreds of sculptures. The sculptures are surprising because they are all made from rubbish, like plastic jewellery, bottles, broken pots, sinks, toilets, and pieces of electrical equipment.

The garden has a surprising history too because it was built in secret! A man called Nek Chand started building the garden in 1957. He chose some land in a forest near the city, where there were rocks and a river. Nobody was allowed to build here, so Nek knew that people wouldn't see him there. Then he started to collect rubbish and make the sculptures.

Nek's project was illegal, but for 18 years, nobody found it. However, in 1975, some government officials discovered it. By that time, the garden was nearly 49,000m² in size. The government tried to destroy the garden, but when the people in the city heard about it, they wanted to keep it.

In 1976 it became a public park. After that, Nek Chand received a salary so that he could work in the garden full-time, and fifty more people were employed to work there with him. People paid to visit the garden. Nek put collection centres around the town so that he could make more recycled sculptures.

However, the garden wasn't safe yet. In 1990, the local government tried to build a road through the park, just for VIPs. But local people protested, and their plan failed. Later, Nek became famous. He travelled to the USA to tell people about his garden. But when he returned, his garden was destroyed! It took a long time to rebuild it and keep it safe.

Now, a charity manages the garden. Five thousand visitors visit this special place each day. Volunteers, workers, and local children still work here. They keep the garden clean and learn to make sculptures. If you like, you can volunteer here too!

Home assignment: Write your MCQs to test understanding of the short texts about teenagers' embarrassing moments.

(1) Lucinda, Huddersfield: "I was sitting down cross-legged in assembly because there was an awards ceremony going on. I was relieved when it was finally over but when I stood up I had really bad pins and needles in my leg and it felt really numb. I tried to walk but my jelly leg just gave way and I tumbled to the floor! Everyone started laughing and I kept getting asked if I enjoyed the trip!"

(2) Hannah, Reading: "I was out shopping with my mum and my sister and as they were paying for their stuff, I asked my mum if I could go and wait in the car. She passed me the keys and off I went. I walked towards the door that said EXIT on it and I thought it was an automatic door. However, I hadn't noticed the sign saying that the door was out of order and to use the one next to it! I walked straight into it and nearly knocked myself out! My mum and sister as well as a couple of staff members came rushing over but my embarrassment made me forget about the pain!"

(3) Rhiannon, Plymouth: "I was at my friend's house and she let me hold her cute new hamster. Suddenly it wriggled out my hand and ran across the top of his cage but it ran too far and fell into the fish tank! My friend got him out and luckily he was OK. I felt really guilty and mega-embarrassed but thankfully my friend forgave me!"

(4) Davina, London: "It was my 13th birthday and it certainly was unlucky for me! My parents thought they would try and be funny and put an old baby photo of me in the local paper announcing it was my birthday. I had a really silly top on and a little ponytail coming out of the top of my head! Unfortunately, someone in my class spotted it and brought it into school to show everyone. I was mega-embarrassed and made my parents swear they would never do anything like that again!"

Performance test: Write your MCQs to test understanding of texts (1-5) below.

1. Ben, Cardiff: "I was at a quiet library, trying to study for my exams. I had brought a flask of soup for lunch. When I tried to open it, the lid was stuck. I twisted it harder and harder, and suddenly it popped open with a huge bang, splashing warm tomato soup all over a girl sitting opposite me and her open laptop. She just stared at me, and I wanted the ground to swallow me whole."

2. Olivia, Bristol: "My family went to a fancy restaurant for my grandfather's birthday. I needed to go to the restroom and walked towards a door that looked like the right one. I pushed it open confidently, but it wasn't the restroom. It was the kitchen, and I walked right into a busy chef holding a massive tray of desserts. He stumbled but thankfully didn't drop anything. All the kitchen staff stopped and stared at me. I just mumbled 'sorry' and ran out."

3. Leo, Glasgow: "I was on a crowded bus, listening to music on my headphones. I was really into a new song and started singing along quite loudly, thinking no one could hear me. After a minute, I noticed an old lady next to me tapping my shoulder. I took off my headphones, and she said, 'You have a lovely voice, dear, but your headphones aren't plugged in'. The whole back of the bus was laughing."

4. Grace, Leeds: "For a school project, we had to give a presentation about a historical figure. I was really nervous. When it was my turn, I walked to the front of the class, but my shoe caught on a cable for the projector. I tripped and fell, sending my carefully organized note cards flying everywhere. My classmates tried not to laugh, but it was a disaster trying to put them back in order."

5. Sam, Dublin: "I was trying to impress a girl I liked by showing her I was good at skateboarding. I tried to do a simple trick, but my board slipped from under my feet. It shot across the pavement and rolled right into a big puddle, splashing muddy water all over a man walking his small white dog. The dog started barking, the man started shouting, and the girl just walked away."

Results of performance test

The test tasks developed by students were assessed against the 10-point holistic scale, whose descriptors were based on the

rules for writing good MCQs (see the scale in Appendix). The scores on the performance test are presented in Table 1.

Table 1

Scores for the performance test

No	Student	Score	Mean	Median
1	V.D.	9.5	8.63	8.75
2	K.B.	8.5		
3	L.S.	9		
4	M.D.	8.5		
5	M.P.	8.5		
6	P.A.	7		
7	T.S.	9		
8	S.L.	9		

The analysis of student scores indicates a generally high level of academic performance within the group. The mean score of 8.625 reflects a strong overall average, while the median score of 8.75 shows that at least half of the students achieved results close to the upper end of the grading scale. The slightly higher median compared to the mean suggests a minor influence of lower scores on the average, indicating a small left skew in the distribution. Overall, the results demonstrate consistent and strong performance across the group, with most students scoring nearer to 9 than to 8.

Overall, the data confirmed that the intervention was effective and, therefore, Research Question 1 was resolved.

Feedback questionnaires

1. Self-Reported Learning Outcomes

This section of the questionnaire presents a checklist of ten "I have learned..." statements formatted as a Likert scale of five options ("1" indicating total disagreement and "5" – total agreement). The checklist is shown in Table 2.

Figure 1 below presents the distribution of ratings for each of the ten statements, based on the Google Forms data; no scores of 1 or 2 were recorded.

As is seen in the graph, the respondents demonstrated robust self-reported progress, with the majority of ratings at the upper end of the scale. All respondents indicated agreement with the statements related to developing MCQs with four answer options

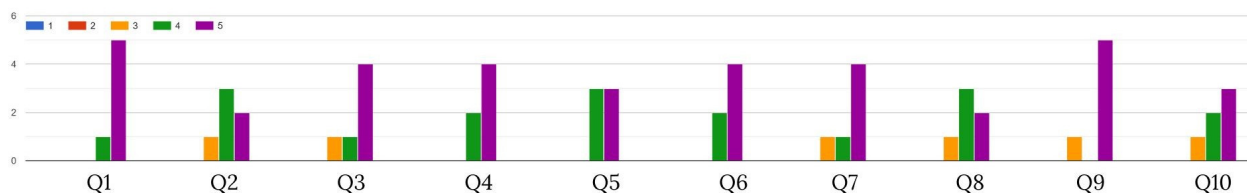
(Q1), formulating questions consistent with the complexity of the texts (Q4), creating grammatically unified options (Q5), ensuring text-dependent answers (Q6), formulating clear stems and distractors in English (Q8), and identifying errors in sample tasks (Q9). For these items, the majority of ratings were 5, with only a few 4s and no 3s. Slightly more varied responses appear for statements

concerning plausibility of distractors (Q2), option lengths (3), avoiding qualifiers (Q7), and ability to develop good MCQs independently (Q10), where a small number of neutral ratings (3) occurred. Overall, 92% of all responses intervention's were at level 4 or 5, underscoring the success in building procedural knowledge and assessment literacy.

Table 2
Checklist of "I have learned..." statements

I have learned to...	
1	develop MCQs with four answer options (one correct, three plausible distractors)
2	develop plausible distractors so that only one answer is absolutely correct
3	develop answer options of similar length and grammatical structure
4	formulate questions so that they are not more complex than the language of the text
5	create grammatically unified answer options (same part of speech or syntactic structure)
6	ensure the correct answer can be determined only from the text, not from background knowledge
7	avoid words and expressions such as "mostly", "completely", "none of the above" in answer options
8	formulate a clear, unambiguous stems and distractors in English
9	identify typical errors in ready-made MCQ tasks and explain their reason
10	independently construct good MCQs

Figure 1
Distribution of Student Self-Reported Learning Outcomes



2. Responses to Open-ended Questions

Question 1. Що саме на занятті – пояснення викладача та/або виконання певної вправи – допомогло Вам найкраще зрозуміти, як створювати запитання з множинним вибором? (Which stage of a class (teacher explanation or performing some task) helped you best understand how to create multiple choice questions?)

A consensus emerged around error-analysis tasks and guided exemplars. Students frequently cited the comparative review of flawed versus revised MCQs as pivotal. One student noted: "When we analysed tasks with mistakes, whether in the structure of the question itself or inaccurate or ambiguous answers, it helped me clearly understand the topic and realise how to correctly develop such tasks and

what mistakes to avoid". Another student commented: "The criteria offered by the teacher (adapted from Christine Coombe – O. K. & Y. K.) were the basis for understanding what makes a correct MCQ. Therefore, it was convenient to edit faulty item using them". The third respondent emphasised as significant "the moment when I learned that answer options should be of the same length and grammatical structure and should focus on understanding the text rather than background knowledge". Task 2 in the intervention was unanimously found meaningful, e.g., "Task 2 clearly showed the structure for this task type".

Question 2. Що Вам було найскладніше під час створення власних завдань на множинний вибір? (What was the

biggest difficulty for you when creating your own MCQs?)

Students' responses aligned with observed post-intervention errors, primarily revolving around distractor design and uniformity. The most common difficulty was maintaining equal length among options, mentioned by several participants: *"It was difficult to create options of the same length"*, *"For me, the hardest point was developing all answer options of the same length"*. Creating plausible distractors was also noted: *"Writing different options, but at the same time making sure they weren't obviously inappropriate"*. A further challenge was ensuring grammatical consistency: *"My personal challenge was making all answer options unified"*. One student summarised: *"Creating distractors according to specific criteria was difficult indeed"*.

Question 3. Які поради Крістін Кум (Coombe et al., 2007) виявилися для Вас найкориснішими у цьому воркшопі? (Which tips from Coombe et al. (2007) proved most useful for you in this workshop?)

Students identified a range of principles as particularly valuable. Equal option length was cited by multiple respondents to prevent test-taking bias toward the longest option. One student highlighted: *"All options should be approximately the same length, as students tend to pick the longest as correct"*. Another respondent recapped the distractor profile: *"In general, among the four options, only one is absolutely correct, the second is absolutely incorrect, the third is correct in terms of reflecting the general meaning, and the fourth is also correct but in too narrow a sense for the given context"*. Others frequently mentioned principles included: avoiding verbatim text repetitions and qualifiers like "mostly", "completely", or "none of the above"; ensuring that questions are based on the text rather than on background knowledge; moving repeated words or phrases from options to the stem; maintaining grammatical parallelism across all options.

Question 4. Які частини воркшопу, на вашу думку, можна було б покращити або доповнити? (In your opinion, which part/s of the workshop could be improved or supplemented?)

Suggestions for improvement were minimal, with most students deeming the workshop comprehensive. Representative comments included: *"In my opinion, everything was clear and understandable"*; *"All parts were complete and instructive for me. Therefore, I think they do not need improvement"*; *"Generally, everything is clear: there are criteria for task development, editing tasks, and original creations"*; *"All the material was accessible and easy to understand, so I think everything was great"*. A few responses proposed minor changes, such as more translation practice for rubrics from Ukrainian to English or additional examples of narrow-context distractors, but overall satisfaction was high.

Question 5. Як ви плануєте застосовувати набуті знання та вміння з теми у своїй подальшій педагогічній діяльності? (How do you plan to apply the acquired knowledge and skills on this topic in your future professional activity?)

The responses revealed strong transfer intentions, with students emphasising assessment design in their teaching practice. One student stated: *"This experience was useful for me. I will definitely apply the acquired knowledge in the future when developing tests, tasks, and assessments for my students"*. Another student noted: *"Without developing multiple-choice tasks, it is impossible to accurately assess students on tests or independent work, so this knowledge will definitely be useful when working in school"*. Another respondent reflected: *"In my pedagogical practice, I will definitely apply the acquired knowledge, especially Christine Coombe's advice. Future practice will enrich my experience in preparing tasks for student assessment"*. One student, currently working as an English tutor, commented: *"I will definitely*

apply this knowledge to create my own tests for school children".

In conclusion, the feedback questionnaire confirmed the effectiveness of the intervention from the students' perspective. The detailed distribution of responses in Section 1 shows that while most principles were strongly endorsed (scores of 5), a few areas such as creating plausible distractors and formulating clear questions received more 4s and an occasional 3, indicating room for further refinement. Nevertheless, 92% of responses were at level 4 or 5, demonstrating that students perceived significant progress in their understanding of MCQ construction principles. The alignment between perceived challenges (option length, grammatical parallelism, creating plausible distractors) and actual error patterns observed in the performance test validates the preventive design of the intervention. Students' strong intentions to apply these skills in their future teaching practice, coupled with their ability to articulate specific principles from Coombe et al. (2007), indicate that the intervention successfully enhanced both their assessment literacy and their metacognitive awareness of test design quality. The use of the native language (Ukrainian) for the questionnaire allowed students to express their reflections fully and authentically, contributing to the richness of the qualitative data collected. To sum up, evidence from the questionnaire allow to consider Research Question 2 in this research study resolved.

Conclusion

The preventive intervention in the LTA course demonstrated its effectiveness in terms of developing students' skills of writing good MCQs, although the absence of the control group belongs to limitations of the study. The impact of the intervention is vividly demonstrated through the quality of the constructed MCQs. All students but one achieved high level of performance rooted

in their in-depth understanding of rules for constructing MCQs to assess reading, demonstrated capacity to apply them in the pre-service context and willingness to transfer the developed skills in future assessment-related situations. The data obtained through a performance test were supported by responses to two sections of the feedback questionnaire, with the open-ended questions providing personal perceptions of the intervention impact on learning outcomes and outlining prospects of applying LTA skills on the job.

Author contributions: O. Kvasova – conceptualization of the study, methodological supervision, development of research instruments and instructional materials, interpretation of findings; Y. Kaleniuk – design and implementation of the preventive intervention, development of research instruments and instructional materials, data collection and analysis.

References

- Adams, T., Koster, B., & den Brok, P. (2023). The effects of an intervention in student teachers' classroom management learning processes. *Pedagogische Studiën*, 101, 60-84. <https://doi.org/10.59302/ps.v101i1.18784>
- Alfonso, M., Busso, M., Ñopo, H., Rivera, A., & Yentzen, T. (2025) Becoming a teacher: Experimental evidence from an information intervention. *International Journal of Educational Development*, 118, 103362. <https://doi.org/10.1016/j.ijedudev.2025.103362>
- Barrio, B. L., Lindo, E. J., Combes, B. H., & Hovey, K. A. (2015). Ten years of response to intervention: Implications for general education teacher preparation programs. *Action in Teacher Education*, 37(2), 190–204. <https://doi.org/10.1080/01626620.2015.1004603>
- Coombe C., Folse, K., & Hublely, N. (2007) *A practical guide to assessing English language learners*. The University of Michigan Press, 2007. <https://doi.org/10.3998/mpub.170430>
- Desimone, L. M. (2009). Improving impact studies of teachers' professional development: toward better conceptualizations and measures. *Educational Researcher*, 38(3), 181–199. <https://doi.org/10.3102/0013189X08331140>.
- Fröjdendahl, B. (2018). Pre- and in-service teachers' assessment literacy – a qualitative approach. *Literacy Information and Computer Education Journal (LICEJ)*, 9, 2886–2894.

- Giraldo, F., & Murcia, D. (2019). Language assessment literacy and the professional development of pre-service language teachers. *Colombian Applied Linguistics Journal*, 21, 243-259. <https://doi.org/10.14483/22487085.14514>
- Kalaizti, C., & Theocharis, D. (2023). A creative writing intervention for second language acquisition development in multilingual students. *Journal of Education and Training Studies*, 12(10), 35-43. <https://doi.org/10.11114/jets.v12i10.6375>
- Kalinowski, E., Gronostaj, A., & Vock, M. (2019). Effective professional development for teachers to foster students' academic language proficiency across the curriculum: A systematic review. *AERA Open*, 5(1), 1–23. <https://doi.org/10.1177/233285841982869>
- König, J., Großschedl, J. & Jost, J. (2025). Effectiveness of interventions in teacher education – curricular and methodological innovations. *Z Erziehungswiss*, 28, 1337–1344. <https://doi.org/10.1007/s11618-025-01355-7>
- Koster, M., & Bouwer, R. (2018). Describing multi-faceted writing interventions: From design principles for the focus and mode of instruction to student and teacher activities. *Journal of Writing Research*, 10(2), 189-224. <https://doi.org/10.17239/jowr-2018.10.01.03>
- Kvasova, O. (2020). Will a boom lead to a bloom, or how to secure a launch of assessment literacy in Ukraine. In D. Tsagari (Ed.), *Language assessment: From theory to practice* (pp. 116-136). Cambridge Scholars Publishing.
- Kvasova, O. (2022a). The impact of student teachers' pre-existing conceptions of assessment on the development of language assessment literacy within an LTA Course. *Languages*, 7(62). <https://doi.org/10.3390/languages7010062>
- Kvasova, O. (2022b). *Workshops on language testing and assessment*. Lira-K Publishing.
- Malone, M. E. (2017, July 17–21). *Including student perspectives in language assessment literacy* [Paper presentation]. Proceedings of the 39th Language Testing Research Colloquium, Bogotá, Colombia (p. 83). <https://www.tirfonline.org/resource/39th-language-testing-research-colloquium-ltrc/>
- Rossi, O. (2021). The role of linguistic knowledge in item writers' assessment literacy and professional practice. [Paper presentation]. Language Testing Forum of UKALTA, Lancaster, UK. <https://www.researchgate.net/publication/356791262>
- Rossi, O., Brunfaut, T., & Harding, L. (2025). What makes item-writing training useful? Learning from trainee perceptions on an online item writing course for language testing. *Studies in Language Assessment*, 14(1), 212-246. <https://doi.org/10.58379/DIJZ6182>
- Sato, M. & Lara, P. (2019). Interaction vision intervention to increase second language motivation: A classroom study. In M. Sato & S. Loewen (Eds.), *Evidence-based second language pedagogy: A collection of Instructed Second Language Acquisition studies* (pp. 1-23). Routledge. <https://doi.org/10.4324/9781351190558-13>
- Scarino, A. (2013). Language assessment literacy as self-awareness: Understanding the role of interpretation in assessment and in teacher learning. *Language Testing*, 3, 309–27.
- Stranovská, E., & Ficzer, A. (2022). *The application of intervention programmes and the development of reading comprehension predictors*. Verlag Dr. Kovac GmbH.
- Stranovská, E., Haringa, I., & Ficzer, A. (2025). The importance of receptive skills intervention in foreign language teaching. *TEM Journal*, 14(4), 3865-3873. <https://doi.org/10.18421/TEM144-89>
- Tsagari, D. (2017, July 17–21). *The importance of contextualizing language assessment literacy* [Paper presentation]. Proceedings of the 39th Language Testing Research Colloquium (pp. 82–83). Bogotá, Colombia. <https://www.tirfonline.org/resource/39th-language-testing-research-colloquium-ltrc/>
- Tsagari, D., Vogt, K., Froelich, V., Csépes, I., Fekete, A., Green A., Hamp-Lyons, L., Sifakis, N., & Kordia, S. (2018). *Handbook of Assessment for Language Teachers*. <http://taleproject.eu/>
- Ukrayinska, O. (2018). Developing student teachers' classroom assessment literacy: The Ukrainian context. In S. Hidri (Ed.), *Revisiting the Assessment of Second Language Abilities: From Theory to Practice. Second Language Learning and Teaching* (pp. 351–71). Springer International Publishing.
- Ukrayinska, O. (2025) Strategies mitigating linguistic challenges in item writing and marking: insights from Ukrainian pre-service teacher training. In O. Kvasova, T. Kavytska & V. Osidak (Eds.), *Transforming Language Teacher Assessment Literacy* (pp. 278-286). Kyiv University Publishing.
- Volante, L., & Fazio, X. (2007). Exploring teacher candidates' assessment literacy: Implications for teacher education reform and professional development. *Canadian Journal of Education*, 30(3), 749-70. <https://doi.org/10.2307/20466661>
- Wohlfart, O., & Wagner, I. (2024) Empowering future educators: Tailored interventions and digital competency development in teacher education. *Zeitschrift für Hochschulentwicklung*, 19(1), 149-170. <https://doi.org/10.21240/zfhe/19-01/08>

Отримано редакцією журналу / Received: 20.03.26
Прорецензовано / Reviewed: 20.04.26
Схвалено до друку / Accepted: 22.04.26

Appendix

Holistic rating scale to assess the MCQs

Score	Descriptor
10	The student-constructed MCQs fully comply with the rules listed below the scale.
9	The MCQs combine the features of 10 and 8.
8	The student-constructed MCQs almost fully comply with the rules listed below the scale
7	The MCQs combine the features of 8 and 6
6	The student-constructed MCQs violate some of the rules listed below the scale
5	The MCQs combine the features of 6 and 4.
4	The student-constructed MCQs violate the majority of the rules listed below the scale
0-3	No assessable MCQs.

Rules for writing good MCQs (after Coombe et al., 2007)

MCQs to assess reading skills have four options, one of which is absolutely correct, and the other three are plausible; the options are of approximately the same length; the language of the questions and options is unambiguous, clear and of the same level of language complexity; all options have a uniform grammatical form and parallel structure; the answers to MCQs reflect the text content rather than the background knowledge on the topic; the items are free from verbatim repetition of chunks of the input text; the distractors do not give away the correct answer; such qualifiers as "mostly", "completely", "almost", etc. are not used; the stems and options are written in grammatically correct English, with proper choice of words and compliance with the level of formality, etc.

Ольга Квасова (Україна)

ORCID ID: 0000-0002-1479-0811

Юлія Каленюк (Україна)

ORCID ID: 0009-0001-7506-4194

**ПРЕВЕНТИВНЕ ВТРУЧАННЯ ДО КУРСУ
"КОНТРОЛЬ ТА ОЦІНЮВАННЯ РЕЗУЛЬТАТІВ НАВЧАННЯ":
ФОРМУВАННЯ УМІНЬ РОЗРОБЛЕННЯ ТЕСТОВИХ ЗАВДАНЬ
ТИПУ "МНОЖИННИЙ ВИБІР"**

Вступ. Сучасна підготовка вчителя іноземної мови передбачає не лише високий рівень володіння мовою, а й розвинену оцінювальну грамотність (*language assessment literacy*). Одним із найбільш складних аспектів для студентів є розробка якісних тестових завдань, зокрема на множинний вибір (*multiple choice questions*), які вимагають дотримання суворих лінгвістичних та технічних принципів розробки. Попри тривалу практику викладання курсу "Контроль та оцінювання результатів навчання", у студентських роботах спостерігається стійка тенденція до порушення принципів створення дистракторів та недотримання паралелізму. У статті описано методику та результати впровадження превентивного втручання, спрямованого на подолання цих викликів.

Методологія. Учасниками дослідження стали 8 студенток третього курсу, які навчаються в Київському національному університеті імені Тараса Шевченка та здобувають кваліфікацію вчителя української мови та літератури і англійської мови. Робота базувалася на змішаному підході (*mixed-methods approach*). Для оцінки ефективності втручання було використано результати підсумкового тесту на розробку MCQs, а також онлайн-опитувальник, що складався з десяти тверджень за шкалою Лікерта та п'яти запитань, які вимагали розгорнутої відповіді. Таким чином, до методів дослідження увійшли кількісний та якісний інструменти.

Результати та обговорення. Аналіз результатів підсумкового тесту підтвердив високу ефективність превентивного втручання: середній бал студентів склав 8.63 з 10, що свідчить про успішне опанування принципів конструювання MCQ. Дані опитування показали, що 92% відауків відповідають високому рівню задоволеності (4 та 5 за шкалою Лікерта). Студенти зазначили, що найбільш корисними були вправи на порівняння правильних і помилкових варіантів відповідей та засвоєння правил щодо однорідності дистракторів. Водночас якісний аналіз відповідей виявив, що найбільш трудомістким аспектом залишається вирівнювання довжини варіантів відповідей та створення правдоподібних, але однозначно

неправильних дистракторів. Більшість респондентів висловила намір застосовувати здобуті знання в майбутній педагогічній діяльності для забезпечення валідності оцінювання.

Висновки. Результати дослідження доводять, що використання цілеспрямованого превентивного втручання, заснованого на аналізі помилок, значно підвищує рівень професійної грамотності майбутніх учителів у області тестування. Матеріали втручання можуть бути інтегровані в курс "Контроль та оцінювання результатів навчання", використані у розробці інструментів оцінювання під час підготовки магістерських робіт, а також для самостійного професійного розвитку.

Ключові слова: превентивне втручання, оцінювальна грамотність (*language assessment literacy*), мовне тестування, завдання з множинним вибором (*multiple choice question*), підготовка вчителів.

BIOS

Olga Kvasova, PhD in Language Education, is an Associate Professor in the Department of Methodology of Teaching Ukrainian and Foreign Languages and Literatures at Taras Shevchenko National University of Kyiv, Ukraine. She is the founding president of the Ukrainian Association for Language Testing and Assessment (UALTA). Her research interests include English language teaching methodology, language testing and assessment, cross-cultural communication, and academic English.

E-mail: olga.kvasova1610@gmail.com

Yuliia Kaleniuk, student of the master's program "Theory of education and teaching methodology of Ukrainian language, literature, and a foreign language in professionally oriented school" at Taras Shevchenko National University of Kyiv, Ukraine. Her academic interests focus on STREAM-based learning, language assessment, and the development of students' communicative competence

E-mail: kalenyuk.yuliya@gmail.com

Автори заявляють про відсутність конфлікту інтересів. Спонсори не брали участі в дослідженні, в аналізі й інтерпретації даних, в описі результатів дослідження, у схваленні публікації до друку.

The authors had declared no conflicts of interest. Sponsors were not involved in the research, data analysis, interpretation, writing, or publication approval process.