O. G. Kvasova
Taras Shevchenko National University of Kyiv

Constructing valid classroom reading test tasks (multiple choice questions) for university students (notes on a workshop)

The paper contains the materials of a workshop aimed to train teachers in constructing and validating reading test tasks in university classroom. Initially, it offers teachers to check their assessment literacy by doing a test. The author further presents brief theoretical input on types of validity relevant for classroom based assessment, alternates it with tasks for teachers to do independently/in pairs/teams, offers tips to construct valid multiple choice questions and takeaway materials to guide teachers in their classroom practice of assessing students' reading skills. List of recommended reading and keys to tasks are provided.
Key words: *construction and validation reading test tasks, multiple choice questions, university classroom, types of validity, classroom based assessment.*

Квасова О. Г. *Конструювання валідних тестових завдань типу «множинний вибір» для перевірки вмінь читання студентів університетів*

У статті представлено матеріали практичного семінару зі створення валідних тестів з читання для студентів вищої школи. На початку вчителям запропоновано виконати тест на встановлення власного рівня компетентності у тестовому контролі умінь читання. Стислий виклад теоретичних засад розробки тестових завдань, зокрема різних типів валідності, що є релевантними щоденній практиці навчання, перемежається із завданнями для виконання в індивідуальному/парному/груповому режимах, до яких подано ключі. Автор також пропонує практичні поради та рекомендації, якими можна користуватися під час підготовки тестових завдань типу «множинний вибір».
Ключові слова: *конструювання та валідація тестових завдань з читання, «множинний вибір». студенти університету, типи валідності, поточний контроль.*

Квасова О. Г. *Разработка валидных тестовых заданий типа «множественный выбор» для проверки умений чтения студентов университетов*

В статье представлены материалы практического семинара по разработке и валидации тестовых заданий по чтению. Учителям предложен тест для проверки своей компетентности в области тестирования. Затем следует краткое изложение теоретических основ тестологии, в частности, освещаются разные типы валидности, имеющие непосредственное отношение к поточному контролю в студенческой аудитории. Подача теории чередуется с практическими заданиями для выполнения индивидуально либо в парах/группах, задания снабжены ключами. Автор также предлагает практические рекомендации и дополнительные материалы, которые можно использовать при разработке заданий типа «множественный выбор».
Ключевые слова: *разработка и валидация тестовых заданий по чтению, «множественный выбор», студенты университета, типы валидности, поточный контроль.*

With the introduction of ECTS in tertiary schools in Ukraine, the teachers of foreign languages are required to write and administer tests as summative assessment more and more frequently. Two crucial questions arise in this connection: how well are Ukrainian university teachers of English prepared for constructing and administering such tests? Can tests written by practitioners with no proper training in the field of language testing and assessment serve a valid tool to measure students' achievements in learning?

The surveys of Ukrainian university teachers' assessment literacy carried out in 2010and 2013 [3] showed that 1) mainstream teachers predominantly write test tasks to assess language-in-use (lexis and grammar) and reading, 2) the most frequently employed testing format is multiple choice questions MCQ, and 3) teachers are concerned with the issues of constructing valid and reliable tests for classroom use. These conceptual areas laid the basis for the development of training module "Constructing and validating classroom reading test tasks".The goal of this publication is to provide university teachers of English with the brief theoretical input, major tasks, tips and takeaway materials for further use in the classroom. These constituents of module's Part 1("Constructing classroom reading test tasks") are presented below.It is desirable that readers do all tasks in consecutive order, check their performance with the keys provided at the end of the article, do additional reading of recommended literature. To begin with, check how testing wise you are.

TASK 1: do a **Test of Language Assessment Literacy**

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

_____ **A** Small-scale pre-testing or trialling
_____ **B** Administering and scoring the test
_____ **C** Familiarising with Specifications and Guidelines
_____ **D** Revising items
_____ **E** Writing items
_____ **F** Constructing test papers
_____ **G** Analysing items statistically
_____ **H** Receiving feedback from pre-testing

2 *Match the definitions (A-H) with the aspects of test usefulness (A-I). There is one extra word. Write the letters next to numbers.*

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

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

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

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

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

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

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

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

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

3 *Match reading skills (17-22) with the formats suitable to test them (A-F). There is one extra format. Write the letters next to numbers.*

17 _____ Finding specific information in a text by scanning. Understanding the main ideas in individual paragraphs or sections of a text.		**A**	Multiple choice
18 _____ Understanding how a text is structured, how it makes sense through its reference backwards and forwards		**B**	Matching headings to texts
19 _____ Understanding main ideas of paragraphs of a text/parts of text		**C**	Sequencing/ ordering
20 _____ Understanding the gist and the details in a text and producing own response		**D**	Matching clauses to gaps in text
21 _____ Understanding a text through understanding the syntax of a sentence and the words around the gaps		**E**	Short-answer questions
22 _____ Accurate comprehension of the text (gist, main and supporting details, inferences, lexis, syntax)		**F**	Matching questions to texts/parts of text
		G	Matching sentences to gaps in text

4 *Match testing formats (23-46) to rubrics (A-I). There is one rubric that you don't need. Write your answers next to numbers.*

23 _____ Matching jumbled questions to parts of an interview.

24 _____ Filling gaps in text with clauses, sentences, parts of text

25 _____ Matching texts and questions

26 _____ Short-answer questions

27 _____ Matching short texts/parts of text and headings

28 _____ Sentence completion

29 _____ Table completion

30 _____ Multiple choice

A Read the text below. Choose the most suitable heading from the list (*A-...*) for each part (*1-...*) of the text. There is one extra heading, which doesn't match any part. Write your answers in the boxes on the separate answer sheet. An example (*0*) has been done for you.

B Read the text below. For questions (1-...) choose the correct answer (A, B, C, or D). Write your answers on the separate answer sheet. An example (*0*) has been done for you.

C Read the text below. Answer questions (*1-...*) below using a maximum of (*three-four*) words. Write your answers in the boxes on the separate answer sheet. An example (*0*) has been done for you.

D Read the text below. Match parts of the text (*A-...*) to questions (*1-...*). There is/are ... extra question(s), which doesn't/don't match any part. Write your answers in the boxes on the separate answer sheet. An example (*0*) has been done for you.

E Read the text below. Some (*number*) sentences have been removed from the text. Choose from sentences (*A-...*) the one which best fits each of (*1-...*). There is one extra sentence, which you do not need to use. Write your answers in the boxes on the separate answer sheet. An example (*0*) has been done for you.

F Read the text below. Complete spaces (1-...)using a maximum of (*three-four*) words. Write your answers on the separate answer sheet. An example (*0*) has been done for you.

G Read an interview with ... . The questions have been mixed up. Match the interviewer's questions (*A-...*) below to answers (*1-...*). There is one extra question, which doesn't match any answer. Write your answers in the boxes on the separate answer sheet. An example (*0*) has been done for you.

H Read the text below. Choose the most suitable summary sentence from the list (*A-...*) for each part (*1-...*) of the text. There is one extra summary sentence, which doesn't match any part. Write your answers in the boxes on the separate answer sheet. An example (*0*) has been done for you.

I Read the text below. Find information in the text to complete boxes (*1-...*) in the table. Write your answers in the boxes on the separate answer sheet using a maximum of (three or four) words. An example (*0*) has been done for you.

(See the answer key at the end of the article)

THEORETICAL INPUT:
Validity is a key test quality

"Validity is the extent to which a test appropriately measures the intended testing area"

[4: 217]. In other words, *does the test measure what it is supposed to measure?*

Today validity is not viewed as something homogeneous but rather as a number of interrelated aspects such as content, concurrent, scoring and face types of validity.

Content validity is based on subject experts' judgements of test content. In other words,*does the content of the test adequately cover all the aspects of language ability we are interested in for making this decision?*

Concurrent validity "involves the comparison of the test scores with some other measures of the same candidates taken at roughly the same time as the test." [1:177]. In other words, *do scores on our test agree with the results of other tests of the same abilities?*

Scoring validity shows the extent to which test scores are based on appropriate criteria; exhibit consensual agreement in their marking; are as free as possible from measurement error;stable over time.In other words,*are the test scores consistent enough for us to have confidence in the results?*

Face validity refers to the "surface credibility or public acceptability"of a test [1:172].In other words, *does test look as though it measures what it is supposed to measure? Is it free from language errors, misprints, poor lay-out?*

Validity is a key quality of test, an"invalid" test cannot be useful as a decision making tool. The process of establishing test validity, or test validation, according to Alderson, Clapham and Wall [1:193], is time-consuming and difficult. Test validation consists in presenting and examining the arguments for using the test as a reasonable justification for taking any decision.

TIPS: How to write a valid reading test for classroom use

- Determine the purpose of testing (type of reading; skills/subskills) at a particular CEFR level
- Determine text characteristics which you will need to test the targeted skills
- Select a suitable text from appropriate sources
- Choose an appropriate testing technique/format
- Write items/tasks
- Trial items/tasks
- Do item analysis using statistics
- Collect feedback from testees
- Review items/tasks

TASK 2: consider the descriptors for reading skills CEFR Level B1 and C1. For each level, single out a) text types and text characteristics,b) reading skills. Underline the key words characterising them.Compare and contrast these characteristics.

Level B1

Straightforward factual texts on subjects related to my field of interest. Everyday material, e.g. letters, brochures and short official documents. Straightforward newspaper articles on familiar subjects and descriptions of events. Clearly written argumentative texts.Personal letters expressing feelings and wishes.Clearly written, straightforward instructions for a piece of equipment. Straightforward factual language, clearly written general argumentation (but not details). Find general information I need in everyday material. Locate specific info by searching one long or several different texts.
Ability to identify main conclusions and follow argument restricted to straightforward texts.

Level C1

Wide range of long, complex texts from social, professional or academic life.Complex instructions on a new unfamiliar machine or procedure outside my area.
Identify fine points of detail including attitudes and opinions which are not explicitly stated. Understand in detail complex texts, including fine points of detail, attitudes and opinions.
Understanding of details of complex texts usually only if difficult sections are re-read.Occasional use of dictionary.

THEORETICAL INPUT:
Types of reading. Reading skills/subskills.

In real-life situations people generally resort to two major **types of reading -** expeditious reading and careful reading.These skills/subskills are targeted at while constructing reading tests.

Expeditious reading includes skills of

A skimming to:
▸ obtain main ideas and discourse topic quickly and efficiently;
▸ establish quickly the structure of a text;
▸ decide the relevance of a text (or part of a text) to their needs
B search reading to quickly find information on a predetermined topic
C scanning to quickly find:
▸ specific words or phrases; specific names
▸ figures, percentages;
▸ specific items in an index

Skills of <u>careful reading</u> include:

A   skills to

▶ identify pronominal reference;

▶ identify discourse markers;

▶ interpret complex sentences;

▶ interpret topic sentences;

▶ outline logical organisation of a text;

▶ outline the development of an argument;

▶ distinguish general statements from examples;

▶ identify implicitly stated main ideas;

▶ recognise writer's intention;

▶ recognise the attitudes and emotions of the writer;

▶ identify addressee or audience for a text;

▶ identify what kind of text is involved (e.g. editorial, diary, etc.);

▶ distinguish fact from opinion/hypothesis from fact/ fact from rumour or hearsay

B   skills to

▶ infer the meaning of an unknown word from context.

▶ make propositional informational inferences, answering questions beginning with *who, when, what*

▶ make propositional explanatory inferences concerned with motivation, cause, consequence and enablement, answering questions beginning with *why, how,* make pragmatic inferences.

TASK 3: complete the following list of **Input text characterisics** that are relevant for teaching and testing reading at the CEFR level you are familiar with:

Text **types**: articles ...
Text **forms**: description ...
Graphic features: tables ...
Length: very short...
Readability (range of vocabulary, grammar): high frequency vocabulary and structure...
Testing formats: multiple choice questions (MCQ) ...

(See possible answers at the end of the article)

BRAINSTORMING:

What testing format appears "by far the commonest way of assessing reading"? Why?

Do you agree that "it is inadequate to measure the uderstanding of text by only one method"? Why/why not?

How far do you agree with the statement "The construction of MCQ is a very skilled and time-consuming business"? Give your reasoning.

What are advantages and disadvantages of writing/administering MCQ for a) teacher testers, b) testee students?

THEORETICAL INPUT:
MCQ structure

A multiple choice question consists of a) a stem and b) responses or options.

A stem can be formulated as a question *Where did John go?* or a statement with a gap to fill *John ... to the store* (see other variations of a stem in further activities).

There are normally **four options** of responses in reading MCQ: one is the correct answer and the rest three are distractors. All distractors should look plausible and, with the exception of the incorrect option, relate to the context.

TIPS:How to write good MCQ

In **stems**, questions/statements should be clearly formulated.The language of the stem and options should not exceed the **difficulty** of the input text. Use of metalanguage is undesirable unless it is the target of testing.

How to construct MCQ options
All options including the incorrect answer should be **plausible**.

A   correct. Unambiguous (**just right**) answer

B   correct but too broad, **general**

C   correct but **too specific**, focuses on one detail

D   incorrect. Looking plausible but the idea is **not developed in the paragraph**.

This pattern is only exemplary. It is a good rule to **rarely** put the correct answer first, however, it should **not** always go third or fourth.

TASK 4 (a):consider the 12 MCQ items below in terms of their a) CEFR level, b) reading subskill tested, c) difficulty. Write out the stems of each item.

Read the texts below and respond to the questions by circling what you think is the best option

1. Peel the potatoes and cut them into small cubes. Boil for about 15 minutes, drain and leave to cool. Mix together with mayonnaise and salt and pepper to taste.

Where is the passage taken from?

a) A menu in a restaurant

b) A recipe in a cook book

c) A restaurant guide

2. Historic industrial buildings in the heart of the city centre have been chosen for possible conversion into a riverside village. Developers are looking to build 80 flats and houses in the old Industrial Conservation Area. Detailed plans for a village-style development have been submitted to the council for approval.

What is the best title for this text?

a) New rural development

b) Approval given for new homes

c) New factory near village

d)   Plans for new housing development

3. **Read the text and answer the question below:**

.......................after strong winds

Last week's storms have left a trail of destruction across the Wye valley and South Shropshire. The worst gales in a decade have left insurance companies facing hundreds of claims as property owners assessed the damage. Hardly a building was left untouched and in Titley one house lost its roof in the 100mph winds. According to Councillor John Price, repairs to Burton village hall are estimated to be in excess of £2,000.

Which words complete the headline (..... after strong winds)?

a) House repaired.
b) Extensive damage.
c) Hundreds homeless.
d) Insurers pay out.

4. **Which of the people mentioned has recently died and is to be remembered in a concert?**

a) Christopher van Kampen
b) Martin Brabbins
c) Alexander Raskatov
d) Oleg Kagan

The concert given by the Nash Ensemble on 5 October as part of the South Bank's Haydn Festival will be dedicated to the memory of Christopher van Kampen. On 13 October, with the viola player Paul Silverthorne and the London Sinfonietta under Martin Brabbins, he was due to give the world premiere of Alexander Raskatov's double concerto 'Misere in memoriam Oleg Kagan'. This concert will now also be dedicated to his memory.

5. A spokeswoman for the RSPCA (the Royal Society for the Prevention of Cruelty to Animals), Jo Crozier, said: "If it hadn't been for DNA testing, we would never have been able to prove that it was these men who killed the animal, even though they were found only 600 yards from the body. All we would have had was circumstantial evidence and that wouldn't have been enough to convict them."

Who or what helped the RSPCA to have the men convicted?

a) Jo Crozier
b) Results of DNA testing
c) Where the men were found
d) Circumstantial evidence

6. **The Holiday of a Lifetime — in China**

A 15 night tour to Beijing, including Xian, Guilin, Shanghai, Suzhou, Wuxi, Guangzhou, Nanjing

From £1,399

The Price Includes:

- scheduled-flights (non-stop) from London/Hong Kong and Beijing/London
- total 15 nights accommodation
- 12 breakfasts, 13 lunches, 11 dinners (including farewell Peking Duck Dinner)
- all excursions/ sightseeing
- all local transport (train/plane/boat)
- 3 evening entertainments
- visa for China
- tour guide
- UK departure tax £12

What is NOT included in the price of this holiday?

a) Food
b) Hotels
c) Air fares
d) Car hire

7. Socialists defended a plan to cut France's work week to 35 hours after President Jacques Chirac criticised the proposal, and a poll said that less than half the French supported it. At the close of a job conference last week, Prime Minister Lionel Jospin said he intended to introduce the 35-hour week by 2000 as a way to create new jobs and reduce the country's high unemployment rate. Many French employers oppose the plan, which they say would increase costs and would not deliver on promises to reduce the 12.5 per cent jobless rate.

Who supports the plan to reduce France's working week?

a) the President
b) most of the people polled
c) the Prime Minister
d) many French employers

8. Suzanne Somers is always going to struggle to lose the image of the superficial character she played in a recent TV drama. Try as she might, she is not particularly convincing here as an ace journalist who sets out to uncover the mystery behind the death of seven people in a night-club. Ed Begley and Michael Nouri offer satisfactory support, but the direction is strictly routine, and the same goes for the script.

What is the writer's opinion of the film?

a) Some of the acting is excellent.
b) The film is well-directed.
c) The leading performance is weak.
d) The script is well-written.

9. The new digital cameras are great fun and very easy to use. They let you review your pictures the moment you take them, so you can re-shoot right away if you are not satisfied. But remember, a digital camera is just a computer **XXXX**. It's not a replacement for your ordinary camera.

What is the meaning of the missing word (XXXX) in the text?

a) "something that is of poor quality"
b) "an item that is not essential, something extra"
c) "something expensive but good value for money"
d) "a fashion which always remains popular"

10. Which sentence best fills the gap in the text?

"I do think this course should be fun," he said. "And people should get out of it what they want. Some come here wanting to be very serious and even a bit competitive about work. And that's fine. Others are very hesitant and find that a bit of humour can relax them.

> a) They tend to learn more.
> b) You get all sorts on these courses.
> c) I prefer the serious student.
> d) We only accept professional people.

I've taught high court judges and dustmen and the mix is part of the fun."

11. Read the text and decide how you would describe it.

With our travel agency, the holiday you book is the holiday you get. If you arrive and find we've failed to live up to our promises, let us know what the problem is within one day of your arrival. We'll spend 24 hours doing everything possible to sort the problem out. In the unlikely event that we can't resolve your problem and make you happy within 24 hours, we'll fly you home and give you your money back.

It is...
a) a commitment
b) an appeal
c) a warning
d) a vow

12. Read the text and then predict the sentence that comes next.

The original cottage gardens were there for only one purpose. They were essential to feed the working man's family. Until the plague in 1348, labourers worked for their masters in return for the rent of their cottages. Mostly they had to be self-sufficient, growing what they could and keeping animals.

a) Therefore, many of them were able to negotiate wages to buy some essentials.
b) Anything they "bought" was obtained by the age-old system of bartering.
c) The plague wiped out about a third of the population, so labour was hard to find.
d) Still, it wasn't much and the cottage garden was always vital to their existence.

(abridged, *from* Ch.J.Alderson's workshop)

TASK 4(b): complete the table with your ideas about test items 1-12 above. Use the following marks:

CEFR Levels: A 1-2; B 1-2; C 1-2
Skill: INF (inference); MI (main idea); D (detail)
Difficulty: low; medium, high
Stem: fill in this grid with all stems

Item	Level	Subskill	Difficulty	Stem
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				

(See suggested answers at the end of the article)

TASK 4 (c): relate the CEFR level, subskill tested and difficulty of each item with input text characteristics. Decide which subskill/s is/are most frequently tested at lower levels and which at higher levels? What input text characteristics prompt testing each particular subskill? What makes some items (very) easy and others (very) difficult? Which items are typical and which are not? Which of the stems would you use in constructing your own items?

TAKEAWAY MATERIALS: in addition to the stems that you have entered in the table in **TASK 4(b),** use stems presented below:

To check understanding of <u>explicitly stated information</u>

▶ According to the paragraph (AtP), which of the following is true of N?
▶ The author's description of N mentions which of the following?
▶ AtP, X occurred because ... .
▶ AtP, N did Y because ...
▶ AcP, why did N did Y?
▶ The author mentions all of the following except ... .
▶ What does the author say about N?

To check the skill to make <u>inferences</u>

▶ Which of the following can be inferred about N?
▶ The author implies that N ... .
▶ Which of the following can be inferred from para X?
▶ It can be inferred from the text that ... .
▶ The text supports which of the following conclusions?

To check the skill of <u>reference</u>

▸ The word "x" in the text refers to ... .
▸ The word/phrase "x" in line 17 refers to ... .
▸ **To check understanding of <u>rhetorical purpose</u>**
▸ The author discusses X in para 2 in order to ...
▸ Why does the author mention X?
▸ The author uses X as an example of ... .
▸ Why does the author use the word "X" when mentioning Y?

To check understanding of essential/key <u>vocabulary</u>

▸ The word "x" in the text is closest to the meaning to ... .
▸ In stating X, the author means that ... .
▸ In line "X", the word "Y" could best be replaced by ... .
▸ **To check understanding of <u>main idea(s)</u>**
▸ What is the author's main point?

TASK 5: to check how MCQ wise you are, consider a reading item presented below. What needs improvement in it?

> According to the text, a telephone-controlled feeder is different than a time-controlled feeder because it can;
> a) Feed the pet when the owner is stuck in the traffic jam.
> b) Provide friendly chat with the pet during its dinner.
> c) Confuse the pet's mind about the dinner time.

(See the answer key at the end of the article)

TAKEAWAY MATERIALS: consider the Common MCQ Item Violations [2: 20-24] to avoid them in your own reading test tasks.

Grammatical inconsistency

Jane spent most of the day at the mall, but she ... anything.

A* didn't buy
B bought
C not buy
D many shops

Extraneous cues or clues

After I've had a big lunch, I only want an ... for dinner.

A pot of soup
B* apple
C big steak
D candy bar

Unparallel options

How do students in the suburbs usually travel to school?

A by bus
B they go by taxi
C *most of them either take the bus or get a lift from their parents
D walk

Giveaway distractors

According to the text, the author of the article comes from ... .

A Dubai
B France
C Buenos Aires
D Disneyland

Double answer or key

The teacher waited in her office until her students ... .

A came
B would come
C come
D had come

TAKEAWAY MATERIALS: use the following **checklist** when writing your own MCQs.

Checklist for MCQ item writing

1. Is the text type worth considering at the targeted CEFR level?
2. Is the text user-friendly in style and acceptable in content?
3. Is the length of text appropriate for your students?
4. Is the level of difficulty appropriate? Are any alterations to the text acceptable in terms of authenticity?
5. Is the MCQ format appropriate for testing the targeted skill/subskill?
6. Is the rubric clear, brief, with limited use of metalanguage?
7. Is the item written in accordance with MCQ specific parameters?
8. Is the number of items acceptable?
9. Do items follow the text sequence?
10. Are the items spread evenly through the text?
11. Can each question be answered independently?
12. Are all items properly labelled: questions with numbers, options with letters?
13. Is the English of text and item grammatically correct?
14. Is the English of stems and options natural and acceptable?
15. Is there just one correct response?
16. Are all the distractors likely to distract?
17. Is there a complete answer key for each task?
18. Is the task lay-out testee-friendly?
19. Is the task economical?

TASK 6: do independent reading of the following **recommended literature**.

Alderson Ch.J. *Assessing Reading*. — Cambridge: Cambridge University Press, 2000. – 398 p.

Alderson, J. C., Clapham C., Wall D. *Language Test Construction and Evaluation*. — Cambridge: Cambridge University Press, 1995. – 310 p.

Coombe Ch., K. Folse, N. Hubley. *A Practical Guide to Assessing English Language Learners*. – The University of Michigan Press, 2007. – 202 p.

Hughes, A. *Testing for Language Teachers.* — Cambridge: Cambridge University Press, 1989. — 264 p.

Into Europe.*Reading and Use of English.*Series editor Ch. J. Alderson.British Council Hungary.Online access: http://www.lancs.ac.uk/fass/projects/examreform/

ANSWER KEYS

TASK 1.1. 1– C, 2 – E, 3 – A, 4 – H, 5 – G, 6 – D, 7 – B, 8 – F
2. 9 – D, 10 – B, 11– H, 12 – I, 13 – G, 14 – E, 15 – C, 16 – F
3. 17 – F, 18 – G, 19 – B, 20 – E, 21 – D, 22 – A.
4. 23 – G, 24 –E, 25 – D, 26 – C, 27 – A, 28 – F, 29 – I, 30 – B

TASK 3. Text **types**: articles, entries, entries, leaflets, letters, forms, diary, maps or plans, fiction, verse, advertisements, post-cards, timetables, reviews, manuals, notices and signs.

Text **forms** : description, exposition, argumentation, instruction, narration.

Graphic features: tables, charts, diagrams, cartoons, illustrations.

Length: very short, longer, lengthy.

Readability (range of vocabulary, grammar): high frequency vocabulary and structure, low frequency vocabulary and structure, idiomatic language, complex syntactical structures

Testing formats:multiple choice questions (MCQ), matching, ordering, gap-filling

Topics, style, intended readership (specific/general), l**ength**

Readability (range of vocabulary, grammar)

Testing formats. MCQ

TASK 4 (b)

Item	Level	Subskill	Difficulty	Stem
1	A 1-2	INF	low	Where is the passage taken from?
2	A2-B1	MI	medium	What is the best title for this text?
3	B1-2	D	medium	Which words complete the headline (..... after strong winds)?
4	B1-2	D	medium	Which of the people mentioned has recently died and is to be remembered in a concert?
5	B1-2	D	medium-high	Who or what helped the RSPCA to have the men convicted?
6	A1-2	D	low	What is NOT included in the price of this holiday?
7	B1-B2	D	med.-high	Who supports the plan to reduce France's working week?
8	B2-C1	INF	med.-high	What is the writer's opinion of the film?
9	B2-C1	INF	high	What is the meaning of the missing word (XXXX) in the text?
10	B2-C1	INF	high	Which sentence best fills the gap in the text?
11	C1	INF	high	Read the text and decide how you would describe it. It is a ...
12	C1	INF	high	Read the text and then predict the sentence that comes next.

TASK 5. 1) There are three instead of four options in this reading MCQ.

2) There are grammar mistakes in the stem and in option c.

3) The stem is badly punctuated.

4) The wording of options is clumsy.

CONCLUSIONS: the workshop presented in the paper was conducted by its author in several universities in Ukraine (Kyiv, Kharkiv, Cherkasy) and highly evaluated by its participants. In their feedback questionnaires the teachers noted that they had received a comprehensive overview of principles lying behind MCQ construction. The enthusiastic teachers received a post-workshop task to construct MCQs of their own teaming up with those working in similar academic context (same syllabus, same course book, same year of study). The syndicates made up of 2-4 members developed one MCQ test task each, tried them out, modified as suggested and further administered the test tasks in their student groups. In many cases teachers administered testee questionnaire to collect feedback on clarity and efficiency of their self-made tasks which helped them determine the tasks' strengths and weaknesses. The data analysed proved the overall effectiveness of the methods of constructing MCQs for classroom use.

The **PROSPECTS** of further research lie in developing methods of constructing valid test tasks of other testing formats (gap fill, short answer questions, etc.) aimed at measuring other reading and/or listening skills.

References

1. Alderson, J. C., Clapham C., Wall D. Language TestConstruction and Evaluation. -Cambridge: Cambridge University Press, 1995. — 310 p.
2. Coombe Ch., K. Folse, N. Hubley. *A Practical Guide to Assessing English Language Learners.*- The University of Michigan Press,2007. — 202 p.
3. KvasovaO., Kavytska T.The Assessment Literacy of University Foreign Language Teachers: A Ukrainian perspective// Language Learning in Higher Education. Volume 4, Issue 1:159–177. DOI: 10.1515/cercles-2014-0010, May 2014.
4. Weir. C. Language Testing and Validation: An Evidence-based Approach. — Palgrave Macmillan, 2005. — 272 p.

Отримано 25.08.2015