Theoretical test

Characteristics

The theoretical test consists of asking the candidate questions or requesting him/her to perform specific tasks in the field of cognitive knowledge and skills, as well as analysing the results of these activities.

Theoretical tests are conducted under controlled conditions. Using this method, many people can be assessed at the same time. The results of theoretical tests are easily related to external standards. At the same time, they are able to be standardised. This means that there is a unified way of using them (including, among others, the order and content of questions, which are usually identical for all candidates). Their relevance (specific learning outcomes have been chosen to be assessed) and reliability (the results are repeatable regardless of time and place) have been confirmed by their use with larger groups of people. Tests should basically measure the quality of knowledge, not its quantity.

The theoretical test, sometimes called a knowledge test, is one of the most commonly used assessment methods. Such a test should not only confirm remembered knowledge, but above all enable the use of this knowledge to be assessed. It allows the use of clear evaluation criteria, while the design of the tools and organisation of the assessment ensure its impartiality and reliability.

In addition, the specificity of the theoretical written test means that the candidate can remain anonymous and his/her personal data can be encoded to increase the objectivity of the assessment. The written test should consist of a large number of test items covering a wide range of issues relating to the given qualification.

Oral tests may (but do not have to) have standardised questions, identical for all candidates. Their advantage is the ability to examine a candidate’s depth of understanding of a given issue.

Theoretical tests can take various forms. They are presented in the table below.

Table 1. Types of theoretical tests

Features	Types of tests
Type of activity performed by the person taking the test*	• Written tests – require written statements from candidates. • Oral tests – require spoken responses from candidates.
Range of use	• Standardised tests (wide use) – used mainly in formal and non-formal education. Expert teams oversee their development, and their use is preceded by a standardisation procedure. This type of test should have instructions for candidates, solution sheets and an assessor's test manual, which includes the technical conditions required to give the test and norms for particular assessment areas. • Non-standardised tests – developed by an assessor’s for his/her own use. They do not undergo a standardisation procedure. The results of such tests cannot be

	compared between individual assessors / test centres.
The purpose of the assessment	• Criterion-referenced tests – their results allow the level of mastery of candidates' learning outcomes to be determined. Their results make it possible to identify the activities / skills which are the most difficult for the examinees and to determine further learning. A special variation of this test is the multistage criterion-referenced test, in which specific groups of test items represent the assessment criteria assigned to individual grades. • Norm-referenced tests – the content of the test items is based on the current results of other people. They are usually used in the selection of candidates for a specific job position or in scientific research.
* This factor is the basis for the categorisation of theoretical tests in the Catalogue of validation methods.	

Tests intended for a greater number of examinees should be standardised. In other words, they should allow the level of learning outcomes to be measured in an objective manner, and their accuracy and reliability should have been previously verified and confirmed in studies using large numbers of people.

The assessor should follow the instructions on how to use the tools to assess the answers.

If open-ended questions are used in the tests, then the answers should be carefully analysed. In case of oral responses, there is a risk that an assessor may not fully convey the information contained in a candidate's statements. This can be prevented, among others, by having the assessor prepare a descriptive assessment, quoting fragments of the answers testifying to the achievement of the given learning outcomes.

Validation stages in which the use of the method is recommended

The reports of Cedefop (an agency supporting the development and promotion of vocational education and lifelong learning in the European Union) show that theoretical tests are mainly used in the assessment stage. This is the case, among others, in the Czech Republic, Estonia, Flanders, France, Portugal, the United Kingdom and South Africa.

Scope of the learning outcomes that may be confirmed using the method

Information on the scope of the learning outcomes that can be confirmed by the theoretical test is for guidance only. It was developed on the basis of analysing the required learning outcomes contained in the Polish Qualifications Framework (second stage descriptors typical of vocational education and training).

It should be remembered that the choice of a given method must always be based on an analysis of the learning outcomes for a given qualification.

Knowledge

The theoretical test may be used to test knowledge, especially factual, at all PQF levels.

Skills

The method is particularly useful for confirming learning outcomes in the areas of:

- information ("calculating, analysing, synthesising and predicting") at PQF levels 2–4 of (relating to retrieving, comparing, assessing and analysing or processing the information needed to perform professional activities);
- at PQF levels 5–6 (relating to analysing the performed profession in consideration of available quantitative data);
- at PQF level 7 (relating to predicting the situation in the profession) and at PQF level 8 (in creating models of development for the profession and drawing conclusions based on them);
- organising work ("planning and revising plans" and "revising activities") at PQF levels 3–4;
- tools and materials at PQF levels 4–6;
- learning and professional development ("supporting the development of others") at PQF level 7.

Social competences

The theoretical test can be used to confirm social competences in the areas of:

- following rules at PQF levels 1–3 (with the exception of providing reliable information on the tasks being performed and compliance with agreements);
- responsibility at PQF levels 2–5 (taking into account the different types of effects of professional activities).

The suggested learning outcomes that can be confirmed using the Theoretical test method

Strengths and weaknesses of the theoretical test

Strengths	Weaknesses
• a commonly known method and judged as accurate and reliable • inexpensive and convenient to use compared to other methods • the results can be easily referenced to external standards • all examinees get the same questions • objectivity of the assessment is maintained (the same criteria) • enables a diversity of questions in formal and substantive areas to be asked; the large number of questions allows many learning outcomes to be covered • easy to assess • ensures the repeatability of the results • the method not very susceptible to the **“examiner's effect”** (the impact of the assessor's subjective assessment of the results,	• not a suitable method to test many skills (especially practical skills) and social competences • ready-made tests cannot be freely adapted to the individual needs of candidates • the structure of the test items can affect the quality of the test (e.g. most items can only confirm reproduced knowledge rather than problem solving) • the way of formulating questions may influence the results (interpretation and answer key) • there is the risk of over-simplification and asking about trivial things • it has a limited number of questions

especially in written tests, can be largely limited)	

Limitations of using the method

The use of the theoretical test is primarily limited by the type of learning outcomes that have been selected for confirmation. This method is mainly intended to assess cognitive knowledge and skills, while the scope of practical skills and social competences that can be assessed with its use is quite limited.

Various factors may have an undesirable influence on the course of the test and final results of the candidate. For example, persons with low verbal and writing skills, persons who dropped out early from the school system, who have trouble solving test problems and those with special needs may have difficulties in demonstrating their learning outcomes with this method.

Required human, organisational and material resources

The required human, organisational and material resources depend on the purpose of the test and the number of persons simultaneously undergoing assessment using this method.

Validation should include the participation of the persons preparing the tests and assessors. The costs of their hiring or training needs to be taken into account.

In order to use the theoretical test method, the following organisational and material resources should be provided:

- preparation of the tests, including their standardisation (if considered appropriate);
- materials required to conduct the test (e.g. sheets of paper, computer stations);
- rooms in which to conduct the test.

The costs of preparing and conducting the test may vary from several hundred PLN (e.g. printing the examination materials) to even tens of millions – it is particularly costly to develop written tests standardised on a large sample.

Awarding bodies should take into account the need to involve additional resources in the case of candidates with hearing or vision disabilities. They must be ensured the proper organisation of the test and adapted materials (e.g. large print question sheets). Conducting the test may require additional skills from both the assessors and candidates (e.g. knowledge of sign language, Braille).

Possibilities of combining the theoretical test with other methods

In Poland, both in formal and non-formal education, theoretical tests are often combined with observation in simulated conditions (practical tasks). In addition, the test may complement or replace unstructured and structured

interviews.

It is also possible to supplement the test in the form of an oral examination with observation in real-life conditions, presentation, unstructured interview and analysis of evidence and statements.

Example techniques

A technique is a way of doing a particular task in a given method, used to collect and analyse data proving that a person has achieved the learning outcomes.

Oral examination

An oral examination usually takes the form of a series of questions posed to the candidates. It allows one to assess, above all, the depth of understanding of a given issue and problem-solving skills. It also allows factual knowledge as well as skills and social competences to be confirmed.

The oral examination is usually not structured, and the assessor adjusts the questions to the candidate's answers on an ongoing basis. This feature makes it similar to the unstructured interview (technique – interview). The interviewer uses interview instructions, i.e. an established objective and loosely formulated topics that can be discussed with the candidate in any way. However, an unstructured interview is not usually used to confirm knowledge, as is the case with an oral examination.

The lack of a specific structure of the oral examination makes it difficult to accurately predict the course of validation using this technique.

The reliability and accuracy of the oral examination depends on the quality and form of the questions (e.g. asking problem questions allows one to confirm the ability to use previously acquired knowledge in new situations, as well as logical thinking). The questions should clearly and precisely assess whether the candidate has mastered the required learning outcomes. They may relate to understanding, justifying or determining causes and effects. The form of the questions should be varied, and the answers should be concise and logical.

It should be remembered that when assessing a candidate's remarks, any interruptions to the conversation (e.g. changing the topic) may affect the final grade. This is particularly harmful when the new thread relates to learning outcomes that are not included in the validation scenario. An assessor who acts in this way is not performing his/her duties properly. However, even in the case of less significant disturbances, the effect of such behaviour can, for example, change the atmosphere of the assessment process. As a result, the candidate may stop responding in detail or honestly. This is also stressful for the person being assessed, who may consequently achieve a worse result.

Several types of oral examinations can be distinguished. Here are three of them:

1. Viva voce

   This is an oral examination consisting of answering questions asked by an examination committee composed of an external and internal assessor. The viva voce technique is mainly used at the university (post-graduate) level after submitting a final work. The objective is to ensure that the candidate has sufficient knowledge about the topic and that the presented work was indeed produced by the candidate.

2. An oral test supplementing observation in simulated conditions (simulation)

This is an oral examination consisting of answering questions asked by the assessor after the candidate has completed a practical task. It is part of the summary session. Usually, there is no formal structure, but the assessors generally ask previously planned questions. It is also possible to ask general questions, which all candidates must answer while performing the practical task.

3. Auditory comprehension test of a language

This test assesses the level of a candidate’s language proficiency. The auditory comprehension test is often performed with devices that reproduce sound files. The test should be matched to the level of proficiency of the examined person in terms of listening comprehension.

Box 1. Stages of oral examination

The course of an oral examination depends on its type, however, the following stages can be distinguished:

1. Introductory question, providing the opportunity to make a spontaneous statement on the selected topic, to analyse an issue.
2. In-depth questions to engage the candidate in a conversation, e.g. "How did you find this out?", "What method did you use to come to this conclusion?".
3. Leading questions, in which the assessor can provide the examined person with some hints, which could lead him/her to the right answer, e.g. "Do you remember the fact xx?", "What is it related to?".
4. Problem-related questions to measure the depth of understanding of issues, e.g. "Can you prove that a given method is better than another one?".

Strengths of the oral examination:

- ensures an immediate result (the assessors receive immediate answers) and enables the candidate's reactions to be directly observed;
- provides the opportunity to formulate additional questions, track the correctness of the thinking process, ask the candidate how he/she arrived at the answer;
- provides the opportunity to become acquainted with the candidate directly and possibly direct his/her further educational path;
- it complements practical tasks well;
- it makes it possible to assess the candidates’ skills of making correct, concise and logical statements;
- it allows assessors to become acquainted with the candidate’s personality traits;
- the candidate has the freedom to formulate answers;
- it prevents cheating.

Weaknesses of the oral examination:

- depending on the number of candidates tested at the same time, the use of this technique can be time-consuming and thus increase the cost of validation;
- clear assessment criteria are required to reduce the influence of the assessor's opinion on the candidate's performance;
- the assessment is susceptible to the influence of subjective and external factors, which can make it difficult to ensure an objective, reliable and accurate assessment;
- the oral examination may be difficult for foreigners who do not know the language used in the assessment;
- in order to increase its reliability, the evaluation of an oral examination should be supplemented with a report on the course of the assessment;
- answers provided during an oral test should be recorded or documented to allow a future appeals procedure or revision of the assessment to be made;
- the length of the oral answer given by the candidate cannot always be predicted;
- there is no repeatability of results;
- there is no possibility to use the results for statistical purposes.

Written test

A written theoretical test is a set of test items to solve, representing a selected range of cognitive learning content. It is usually performed during a single session.

The activities of the person performing the test may include:

- repeating information from memory,
- understanding,
- solving a known problem,
- solving a new problem,
- critically evaluating a situation,
- synthesising knowledge.

The written test uses the test item as a tool. This is a relatively independent, separately scored item whose solution consists of choosing an answer from a proposed list of answers or providing one's own answer.

A case study is a particular type of written theoretical test. This technique can be used during observation in simulated conditions (simulation) and in a written theoretical test.

Box 2. Case Study

The case study has many definitions. The term may refer to research methods used in the social sciences and to didactic methods. For this reason, it can be understood differently by various persons.

With regard to validation, a case study is a technique for confirming learning outcomes, consisting of completing a task or answering questions on the basis of information contained in a description of a specific situation (real or imaginary). The result of the case study can consist of, among others, a diagnosis of a given event, a proposal for solving the problems presented in the case description, a prediction of how the situation may develop further, or a proposed action.

The case study can be used to assess, among others, the ability to analyse information, draw conclusions, think critically and apply knowledge in practice. For example, a candidate may be asked to determine the reasons for a given situation, indicate the main problems raised in the case description, indicate the relationship between the facts, predict or assess the effects of the described activities / events.

Examples of the use of the theoretical test

Foreign examples

Examples of using the theoretical test are found in the Database of Good Practices. Validation, ensuring the quality of validation and certification.

This method is used, among others, in the validation of the following qualifications:

- child care worker in Estonia – written test technique,
- administrative assistant in Italy – oral examination technique.

Polish example

The following example was developed based on the guide for the authors of exercises and test exams prepared by Henryk Rebandel.

Institution Name	**Warsaw Medical University** **Department of Didactics and Learning Effects of the Faculty of Health Sciences, Warsaw Medical University** **Laboratory of Evaluation and Improvement of Medical Education, Office for University Examinations**
Country	Poland
The institutional context of the technique	The Office for University Examinations, established on 22 May 2014, is directly responsible for comprehensive handling the test exams conducted by the Warsaw Medical University The main tasks of the Office include: • handling the paper versions of tests conducted by the scientific and didactic units of the Warsaw Medical University, in particular: ◦ preparing and printing exam papers, ◦ analysing the qualitative parameters of exam questions, ◦ archiving data; • implementing and supporting the University’s electronic examination system, in particular: ◦ coordinating the work of establishing and developing the examination questions database, ◦ analysing the qualitative parameters of exam questions, ◦ conducting electronic exams and supervising their proper course, ◦ verifying the results of electronic exams; • cooperating with the Department of Didactics and Learning Effects as well as with the scientific and didactic units at the Warsaw Medical University in the field of preparing and conducting tests.
Name of the qualification	doctor of medical sciences, doctor of pharmaceutical sciences, first cycle studies diploma in dental techniques, and other qualifications
A brief description of the method used	Theoretical test
Possibilities of acquiring information about the method on one’s own / with the help of a counsellor	At the beginning of a given teaching period (usually at the beginning of the academic year), students are given information about examination requirements: the form, volume and duration of the exam, as well as the threshold criterion for passing it (minimum number of points).
Factors underlying the choice of the method given the circumstances	The test method was chosen because an in-depth examination of students' knowledge can be conducted. This is possible due to, among others: the large number of test items, identical sets of test items for all examinees, short examination time (which is significant in the case of large groups of people), an objective scale of grades (awarding points).

	A well-designed test allows a quick assessment to be made about the level of mastered learning outcomes in a given subject area.
Course	The student can take the test after confirming his/her identity with a personal identity card. The student's entry into the examination room is noted on an attendance sheet, which allows the examiners to determine that all persons taking the exam have submitted their examination sheets. To exclude any mistakes and cheating, the test (examination) sheets are personalised (name, surname, grade book number) and signed by the examinees. Students also sign test books. This allows the examiners to identify persons who did not return the test book after the exam. Test books are treated as a copyright protected publication intended for use for a specific purpose, i.e. only during the test. This is to prevent unauthorised disclosure of the test items. Exam rooms should be large enough to allow adequate space between the persons taking the exam. Students take seats in the room according to a list, usually in alphabetical order, which reduces the probability that people who could "help each other" in answering will sit next to each other. The possibility of cheating is also limited by ensuring an adequate number of people to supervise the course of the test, taking materials from persons who may be trying to cheat and using several versions of the test, so that people sitting next to each other have different items to solve. The persons conducting the test are obliged to inform the students about the applicable rules before distributing the exam papers. They should inform about: the duration of the test, the way the exam will be started, the time and manner of completion, how questions may be asked, how to proceed in the event of errors on the exam papers, how to proceed should the exam be completed earlier, the conditions under which sanitary facilities may be used, rules regarding mobile phones, the consequences of confirmed misbehaviours. The actual starting and ending time of the test is determined after the test books have been distributed to all examinees. Persons conducting the test should inform the examinees how much time is left to complete it several times during its course. After the time set for the test has elapsed, the candidates should remain in their seats until the test books and exam papers are collected. When determining the time needed to solve the test items, all the actions relating to solving them should be taken into account, but the time should not be too long. This allows the test to differentiate between the persons who are well prepared and those who are not. In addition, the test: • can consist of up to 100 questions; • may contain a maximum of 5 response variants (A, B, C, D, E). A maximum of 4 versions of the test can be prepared.
Result of the validation process	The test is used to assess learning outcomes. The test may be passed or failed. Some learning outcomes are not anticipated to be assessed by the test.

Organisational and material resources	To conduct the test, a room is needed that is suitable for the number of examinees, taking into account the need to maintain appropriate space between them. This method requires equipment in the form of computer hardware and software for entering responses and calculating results, as well as testing statistical and psychometric parameters. One should also remember the costs associated with printing the appropriate number of exam books and personalised exam papers in several versions. One should also ensure the appropriate number of people to conduct the test and supervise the examinees. A data bank of ready, standardised test items should be prepared. New items can be prepared for subsequent tests, but it is better to select at least some of the items with known parameters from the data bank, with new items comprising only a certain part of the total (e.g. 20%). It is recommended that test items be prepared by at least several specialists in a given field – this guarantees the quality and diversity of the measurement tools used.
Limitations of use	The test is not always an appropriate method for examining specific practical skills or solving problems in non-typical situations. It takes time and experience for the persons preparing the test items to create a large data bank of them.

Bibliography

American Educational Research Association, American Psychological Association i National Council on Measurement in Education (1999). *Standards for educational and psychological testing.* Washington: AERA. Downloaded from: http://www.apa.org/science/programs/testing/standards.aspx

European Commission, DG Education and Culture i Cedefop (2010). *European inventory on validation of validation of non-formal and informal learning 2010. Thematic Report: Assessment methods*. Luxembourg: Publications Office.

European Commission, Cedefop i ICF International (2014). *European inventory on validation of validation of non-formal and informal learning 2014. Thematic report: validation methods.* Luxembourg: Publications Office.

Jablow, P. (2015). *5 things to know about the financial cost of testing.* Philadelphia Public School. The Notebook. Downloaded from: http://thenotebook.org/latest0/2015/11/24/5-things-to-know-about-the-financial-cost-of-testing

Kubiszyn, T. i Borich, G. (2009). *Educational Testing and Measurement: Classroom Application and Practice.* John Wiley & Sons, Inc.

Announcement of the Director of the Central Examination Board of 1 September 2016 on detailed methods of adjusting the conditions and forms of exams confirming qualifications in the professions to the needs and abilities of candidates with special educational needs in the 2016/2017 school year. Downloaded from: https://cke.gov.pl/images/_KOMUNIKATY/Komunikat%20o%20dostosowaniach%20w%20roku%20szkolnym%202017-2018.pdf

Łobocki, M. (2000). *Metody i techniki badań edukacyjnych [Methods and techniques of educational research]*. Kraków: Oficyna Wydawnicza „Impuls”.

Łoś, E. and Reszka, A. (2009). *Metody nauczania stosowane w kształtowaniu kompetencji kluczowych. Matematyka. Podręcznik metodyki operacyjnej [Teaching methods used in shaping key competences. Mathematics. Manual of operational methodology]*. Lublin: Innovatio Press Wydawnictwo Naukowe Wyższej Szkoły Ekonomii i Innowacji.

Websites:

Central Examination Board: https://www.cke.gov.pl/

Fundacja Vocational Competence Certificate: http://vccsystem.eu/

Rebandel, H. *Poradnik dla autorów zadań i egzaminów testowych [A guide for the authors of exercises and test exams].* Warsaw Medical University; http://biurojakosci.wum.edu.pl/content/poradnik-dla-autorow-zadan-i-egzaminow-testowych-0