

# Multiple-Choice Tests: Objectivity or Delusion in Assessment? A Comparative Analysis from Romania and Moldova

Roza DUMBRAVEANU<sup>1</sup>

Valeria BACIU<sup>2</sup>

Gabriela GROSSECK<sup>3\*</sup>

Daniel-Alex MILENCOVICI<sup>4</sup>

<sup>1</sup> “Ion Creangă” State Pedagogical University of Chișinău, Moldova; <https://orcid.org/0000-0003-2232-3620>; [dumbraveanu.roza@upsc.md](mailto:dumbraveanu.roza@upsc.md)

<sup>2</sup> “Ion Creangă” State Pedagogical University of Chișinău, Moldova; <https://orcid.org/0009-0005-7725-1371>; [baciu.valeria@upsc.md](mailto:baciu.valeria@upsc.md)

<sup>3</sup> West University of Timisoara, Romania; <https://orcid.org/0000-0001-8933-1022>; [gabriela.grosseck@e-uvt.ro](mailto:gabriela.grosseck@e-uvt.ro)

<sup>4</sup> West University of Timisoara, Romania; <https://orcid.org/0009-0001-7536-8884>; [daniel.milencovici@e-uvt.ro](mailto:daniel.milencovici@e-uvt.ro)

\* corresponding author: [gabriela.grosseck@e-uvt.ro](mailto:gabriela.grosseck@e-uvt.ro)

**Abstract:** *Multiple-choice tests (MCTs) are widely used in student assessment due to their perceived objectivity, validity, and reliability. In the digital age, MCTs have become increasingly popular supported by platforms that enable automated grading and real-time feedback. While MCTs offer benefits such as quick grading, minimized evaluator bias, and large-scale scalability, their effectiveness in providing an accurate assessment of student learning, critical thinking, and deep conceptual understanding remains a matter of ongoing scholarly debate. Reliance on testing may result in the oversimplification of knowledge understanding, the encouragement of surface learning, and limitations in assessing higher-order cognitive skills, that raise essential questions about the pedagogical value of MCTs, particularly when used as the main tool for the evaluation of student achievements. The objectives of this paper are to identify the benefits and limitations of MCTs in Higher Education as described in literature and to compare the practices from Romania and Moldova, with a focus on validity, fairness, and impact on student learning outcomes. The study is based on qualitative data collected through surveys and interviews with higher education target representatives in both contexts. Findings highlight common concerns related to guessing the correct answers, the risk of teaching to the test, the misleading design of questions, and weak alignment between learning outcomes and assessment practices. The study suggests remediation of the identified problems by integrating MCTs within a broader, balanced assessment strategy aligned with the principles of constructive alignment to support students' competence development.*

**Keywords:** *multiple-choice tests, digital assessment, higher education, constructive alignment, comparative study (Romania–Moldova).*

**How to cite:** Dumbraveanu, R., Baciu, V., Grosseck, G., & Milencovici, D.-A. (2025). Multiple-choice tests: objectivity or delusion in assessment? A comparative analysis from Romania and Moldova. *Revista Românească pentru Educație Multidimensională*, 17(3), 505-531. <https://doi.org/10.18662/rrem/17.3/1032>

## Introduction

Multiple-choice tests have become a dominant assessment method throughout the assessment services market and education over the past century. These kinds of tests are valued for their standardization, compared to other formats, making them particularly attractive for examinations with large numbers of learners (Bateson & Dardick, 2020). Their widespread use extends to various Higher Education contexts, including university entrance tests, formative assessments, and final-semester examinations.

The popularity of multiple-choice questions in educational settings can be attributed to several reasons. Academic staff and institutions value them for their perceived objectivity and possibility to provide rapid feedback, ease of administration, scoring and reporting, particularly in large classes.

Although there is a wide preference for test-based evaluation methods, MCTs face ongoing criticism. Some researchers argue they may not effectively assess deeper understanding and are vulnerable to guessing, leading to recommendations that they should be avoided in summative assessments (McKenna, 2018). The balance between the practical value of multiple-choice testing and its limitations in assessing complex learning outcomes underpins ongoing debates about its pedagogical relevance in educational assessment (Berg & Singh, 2024).

This article explores the dual nature of multiple-choice testing: its strengths, its limitations, and the potential misconceptions that may arise when such tools are used uncritically and inappropriately, with particular attention to its application in higher education settings in Romania and Moldova. The authors aim to uncover whether MCTs provide a credible insight into learning or an oversimplified approach to assessment.

The objectives of the research are:

- To identify the benefits and limitations of MCTs as described in literature and perceived by teachers and students with a focus on validity, fairness, and impact on student learning outcomes.
- To investigate whether multiple-choice tests are really objective, or whether they risk promoting superficial learning and undermining deeper understanding.
- To compare the practices and attitudes toward the use of MCTs in assessment across higher education institutions in the two countries.
- To provide evidence-based recommendations for improving assessment strategies and integrating multiple-choice tests more effectively within a balanced and meaningful evaluation framework.

The study hypothesizes that when multiple-choice tests are used in isolation, they tend to promote surface learning and fail to adequately assess students' critical thinking and conceptual understanding. Moreover, the lack of alignment between such assessments and intended learning outcomes in higher education undermines their pedagogical effectiveness, particularly in fostering deep learning and competence development.

## **Research methodology**

### ***Research design***

The authors have adopted a qualitative, exploratory research design by combining literature analysis with focus group discussions to critically examine the perceived objectivity and pedagogical value of multiple-choice tests in higher education. The mixed approach enabled both a conceptual and practical understanding of how MCTs function as tools for assessment and how they are perceived by key stakeholders: students and teachers.

A literature analysis was conducted to establish the theoretical and empirical foundations of the article. Sources included:

- Peer-reviewed academic journal articles on assessment strategies.
- Books and articles discussing the design, application, and evaluation of MCTs.
- Institutional policy documents and guidelines from the European and the national universities related to assessment practices.

Documents were retrieved from academic databases including ERIC, Scopus, Web of Science, and Google Scholar, with search keywords such as “multiple-choice tests”, “objectivity in assessment”, “test design”, “assessment in higher education”, “constructive alignment” and “formative/summative assessment”. The inclusion criteria limited sources to English-language publications from 1990 to 2025 that addressed assessment design in higher education institutions.

### ***Focus group discussions***

Two focus group discussions were conducted with university students ( $n = 60$ ) and with academic staff ( $n = 30$ ) representing various departments from several national universities in Romania and Moldova. The participants were selected using purposive sampling to ensure diversity in studied subjects and experience with MCTs. Each focus group session lasted approximately 90 minutes and was moderated using a semi-structured interview guide, which included questions, related to:

- opinions about fairness and objectivity in multiple-choice testing;

- experiences with MCTs as learners (students) and test designers (teachers);
- views on the cognitive demands of MCTs (recall, understanding, problem-solving);
- considerations about the students' workload and questions' design;
- suggestions for improving assessment practices in higher education.

### ***Teachers' selection methodology***

The objective was to select a diverse and representative sample of teachers who are involved in the design, implementation, and evaluation of multiple-choice tests in order to explore their perceptions regarding the objectivity of MCTs in educational assessment. The following criteria were used in the selection process.

- Teaching Experience. Minimum of 5 years of experience in Higher Education.
- Assessment Involvement. Active involvement in student assessment using multiple-choice tests.
- Subject Area. Priority given to teachers from domains where MCTs are commonly used (Psychology, Education, Medicine).
- Geographical Representation. Teachers from both Romania and the Republic of Moldova.

All participants were informed about the voluntary nature of participation, confidentiality of their responses, and their right to withdraw at any time.

### **2.4. Students' selection methodology**

The selection of students was conducted to obtain relevant, diverse, and insightful perspectives regarding the objectivity of multiple-choice tests. It aimed to ensure relevance to the topic, representation across various academic levels and backgrounds, and manageability of group size to allow meaningful participation. Students were selected based on the following criteria.

- Current enrollment in a university-level course that uses multiple-choice tests as an assessment tool.
- Recent experience with MCTs (within the current or previous academic semester).
- Variation in academic results (high, average, and low).
- Representation from different years of study and academic domains.
- Consideration for gender balance and socio-cultural diversity.

Teachers and students were selected through email invitations, and direct contact by the researcher. The recruitment message included a brief

explanation of the study's purpose, the voluntary nature of participation, and assurances of confidentiality and anonymity. Confidentiality was secured through anonymized data reporting, and no identifying information was collected during the sessions.

### ***Data analysis***

The data obtained from the literature review and focus groups were subjected to content analysis. Particular attention was paid to issues like:

- perceived objectivity and reliability of MCTs;
- the alignment of MCTs with learning outcomes and competencies;
- design correctness and sources of bias;
- alternative assessment strategies.

The convergence and divergence between literature-based insights and focus groups' opinions were analyzed to provide a comprehensive and balanced view on the considered issue. The findings are context-specific, based on a limited participant sample, and may not be applicable to all higher education institutions. Moreover, the sample size, while sufficient for qualitative exploration, restricts statistical generalization. Nevertheless, the combination of literature and empirical cases contributes to a better understanding of the perceived strengths and shortcomings of MCTs.

### **Theoretical foundation: literature review**

The existing literature from Romania and Moldova offers limited coverage of the issue of assessment in higher education. Few scholarly resources explicitly address this subject, and comprehensive analyses of digital assessment strategies and methods, particularly regarding the suitability of multiple-choice testing, are scarce. A bibliometric analysis of Romanian Educational Research (RER) in Web of Science was performed recently that examines the evolution and structure of RER through a bibliometric analysis of 10,396 publications indexed in the Web of Science Core Collection from 1975 to 2024 (Cretu & Grosseck, 2025). Five primary research clusters emerged, including studies on student performance factors, digital transformation in education, curriculum development and students' skills, competency-based education and teacher training, as well as advancements in teaching, learning, and assessment practices. This extended research did not reveal any major studies or comprehensive reviews within the Romanian research landscape specifically dedicated to the digital assessment of students' competences in higher education. There is a noticeable gap in the analysis of assessment strategies, methods, and digital tools pedagogically integrated to evaluate learning outcomes and competences at the tertiary level. While some

individual studies touch on related topics, a systematic exploration of digital assessment practices and their effectiveness remains absent in Romanian educational research.

This subject is especially relevant in light of the increasing emphasis on quality assurance and competency-based assessment in European policy frameworks, such as the *Bologna Process* (EHEA, 2020), the *Standards and Guidelines for Quality Assurance in the European Higher Education Area* (ENQA, 2015), and recent policy guidance promoting the pedagogical integration of digital tools to assess learning outcomes and competences in higher education (European Commission, 2023)

A study (Grosbeck et al., 2024) addresses teachers' digital competences, their actual use of digital tools for assessing student learning outcomes, and their training needs. Respondents were grouped into three clusters based on self-rated digital competences. There were clear variations across these groups in how frequently and confidently they applied digital assessment tools in teaching and assessment tasks. Teachers described using digital tests, e-portfolios, peer/self-assessment, gamified elements, and hybrid evaluation formats. However, the study does not deeply analyze or compare specific assessment strategies or methodology frameworks in terms of competence-based outcomes. The study reveals significant obstacles: lack of institutional support, insufficient training, limited AI literacy, and uncertainty around assessment design aligned to competencies.

The focus of this study is on teacher perspective and tools' adoption, not on how different assessment methods measure or impact student competences directly. There's no comparative analysis of assessment strategies, nor deep exploration of validity or effectiveness of used tools or methods in the context of competence-based education. The survey gives valuable insight into teachers' readiness, usage patterns, and training requirements regarding digital assessment in Romanian higher education. It functions more as a needs analysis rather than as an evaluation of how well these digital methods measure student competences. A comprehensive understanding of the impact of diverse assessment strategies on students' learning outcomes and competence development necessitates further investigation, particularly through empirical or experimental research designs.

Educational research in Moldova faces similar challenges to those identified in Romania, particularly regarding the lack of focused studies on digital assessment of students' competences in higher education. One relevant study from Moldova addressed student assessment in higher education, highlighting the pedagogical dimensions of the topic within the framework of a Tempus project (Dumbraveanu, 2006). However, at that time, the digital

dimension was not yet a central concern. Moreover, contemporary educational contexts have witnessed the emergence of new assessment methods, often integrated with digital tools, whose effective implementation requires dedicated investigation and supportive policy frameworks. Today, e-assessment is closely linked to the broader context of e-learning, and its effective implementation requires a coherent and unified strategy. Unfortunately, Moldova does not currently have a national e-learning strategy to guide the integration of digital tools and methodologies into higher education, leaving institutions without a clear framework for adopting and standardizing digital assessment practices (Dumbraveanu & Peca, 2022).

National education strategies in both Romania (Strategia Națională pentru Învățământ Terțiar 2015–2020) (Romania, Ministry of Education and Research, 2015) and Moldova (Strategia de dezvoltare „Educația 2030”) (Republic of Moldova, Government, 2023) align with EU priorities in their stated declarations related to the improvement of assessment systems. However, there is a discrepancy between policy objectives and the limited academic attention paid to assessment approaches and practices in higher education, as well as research priorities; there is a strong necessity for rigorous investigations in this area.

### ***The objectivity and reliability of MCTs***

In Higher Education, multiple-choice tests represent structured assessment tools consisting of a series of questions, each with a stem (the main question) and a set of answer options, typically including one or more correct answers (the key) and several incorrect but reasonable alternatives (distractors). MCTs are used to assess students' knowledge, comprehension, and cognitive processing of course content across a variety of academic disciplines (Rodriguez-Torrealba et al., 2025). These tests are commonly implemented at bachelor level, particularly in large lecture-based courses, e-learning environments, and standardized exams due to their scalability and efficiency in evaluation.

The MCTs have gained widespread acceptance for their perceived objectivity and practical advantages (Butler, 2018). One of the primary reasons MCTs are viewed as objective is their use of standardized scoring procedures, which minimize direct human intervention in grading. Unlike essay-based assessments or oral examinations, where interpretation and personal bias may influence the evaluator's judgment, multiple-choice items have predetermined correct answers. Whether scored manually or via digital platforms, each student's answer is assessed against an identical answer key, ensuring uniform assessment regardless of who marks the test or when it is marked (Haladyna & Rordiguez, 2013). This level of reliability is rarely achievable in other kind of so-called subjective assessments, justifying the use of MCTs as fair and

impartial measurement tools (Nitko & Brookhart, 2014). This standardized scoring seems transparent in academic evaluations; students are less likely to question their results when answers are objectively verifiable, thus reducing disputes and increasing confidence in the fairness of the assessment system.

Multiple-choice tests are also regarded as effective tools for the assessment of lower-order knowledge, such as factual recall, basic understanding, and recognition. These skills form the foundation of knowledge in most academic disciplines, and MCTs can assess them efficiently through well-constructed items. Their structured format makes it possible to include many questions in a short time, thereby enabling broad content coverage across a curriculum (Rodriguez, 2003). This enhances content validity, as teachers can test a range of learning outcomes within a single examination. MCTs facilitate the assessment of specific, discrete portions of knowledge without requiring elaborate explanations from students. The capacity to quickly test and identify gaps in basic knowledge makes MCTs a valuable diagnostic tool early in a course or program (Popham, 2008).

The objectivity of MCTs is justified by their compatibility with automated testing systems, which simplify both administration and scoring. In large university courses, particularly at the undergraduate level, teachers often face the challenge of assessing hundreds of students within tight timelines. MCTs solve this problem by enabling automated grading, often integrated into learning management systems, which produces immediate results and feedback (Pitt & Quinlan, 2022).

Therefore, the perceived objectivity of multiple-choice tests is based in their reliance on standardized scoring, their efficiency in testing factual knowledge, and their suitability for large-scale, automated assessments. These characteristics make MCTs a tempting choice in higher education, especially when reliability, fairness, and resource constraints are the first concern. However, while their objectivity in scoring is largely undisputed, ongoing debates exist regarding their adequacy in measuring complex cognitive processes. For this reason, MCTs are most effective when used as part of a balanced assessment strategy that incorporates both objective and subjective formats to capture the full range of student learning outcomes (Kent-Waters et al., 2018).

### ***The illusion of comprehensive assessment by multiple-choice tests***

Despite MCTs mentioned benefits, there is a growing concern that MCTs offer only a superficial measure of student understanding (Nicol, 2007). The illusion of comprehensive assessment is shattered by their overreliance on recognition, vulnerability to guessing strategies, and inability to assess higher-order thinking skills (Kuechler & Simkin, 2010).

One of the main criticisms of MCTs is that they predominantly assess recognition memory rather than deep comprehension or application of knowledge. In most cases, students are not required to provide information by themselves, but rather to identify correct answers from a list of options. This format allows learners to rely on hints and previous experience, which can lead to correct responses without fully understanding the underlying concepts (Scouller, 1998). For example, the students might successfully identify the definition of a technical term by recalling its phrasing from a textbook or lecture note, even if they are unable to explain it or apply it in a real context. As a result, MCTs may misrepresent the students' actual grasp of the material, creating a false sense of understanding. The depth of learning, very crucial in higher education, is often not captured by such a fixed format (Biggs & Tang, 2011).

Another aspect contributing to the illusion of comprehensive assessment is the influence of guessing strategies. Students' use of test-taking techniques may include the elimination of implausible options, the identification of patterns in answer sequences, and the selection of the most comprehensive answer when uncertainty arises. These strategies, while clever, do not reflect true understanding and can artificially inflate grades (Rios et al., 2022). Research has shown that students with strong test-oriented skills can rank above peers with similar or even greater subject knowledge simply by exploiting the structure of multiple-choice questions (Tarrant & Ware, 2012).

The most significant limitation of MCTs lies in their inappropriateness to adequately assess higher-order thinking skills such as critical analysis, creativity, synthesis, and problem-solving (Goss, 2022). While it is possible to design complex multiple-choice items that require some level of application, these questions are difficult to construct and are rarely used in practice.

Disciplines that focus on interpretation, argumentation, innovation, or ethical reasoning require assessment methods that go beyond the selection of a pre-determined answer. Competences like constructing coherent arguments, analyzing different opinions, or designing original solutions are inherently open-ended and cannot be evaluated through multiple-choice formats alone (Biggs & Tang, 2011).

Even though MCTs are widely used, they offer only a limited and often misleading image of student learning. MCTs create an illusion of comprehensive assessment by focusing on recognition over knowledge application, rewarding guessing strategies, and failing to assess higher-order cognitive skills. For a truly valid and holistic assessment of student achievements, multiple-choice tests should be supplemented with alternative forms of evaluation, such as essays, case studies, portfolios, or oral

examinations that allow for the demonstration of critical thinking, deep understanding, and applied knowledge (Push, 2018). In the context of higher education, where intellectual growth and skill development are crucial, reliance on MCTs alone is insufficient for assessing the full complexity of student learning.

### *The misleading question design*

Research indicates that well-designed MCTs can strengthen retention of information and effectively measure students' basic understanding across various subjects (Douglas & Purzer, 2015). MCTs are often considered efficient tools for measuring a wide range of cognitive skills, from basic recall to higher-order thinking (Oc & Hassen, 2025). Creation of effective MCTs requires careful attention to item design and efficiency of evaluation. The quality of MCTs in actual classrooms varies considerably. A study conducted at a Canadian university (Dibattista & Kurzawa, 2011) examined 1198 multiple-choice items across sixteen classroom tests, revealing that many items and distractors were poorly constructed, indicating considerable potential for improvement in test quality (Slepko et al., 2021). This indicates that, although MCTs offer theoretical advantages, their practical implementation often falls short of ideal standards. Item analysis is essential to ensure validity, reliability, and fairness of the assessment (Alordiah & Oji, 2025).

Misguided question formulations result in MCTs skepticism toward the true value of assessments. Poorly formulated multiple-choice questions can confuse rather than assess. Questions may include ambiguous wording, irrelevant distractors, or double negatives, all of which can distort the student's understanding of what is being asked (Tarrant & Ware, 2012). These design related inadequacies may lead to incorrect responses even from well-prepared students and shift the focus from knowledge assessment to test-taking skills. Moreover, questions that rely heavily on memorization of trivial facts rather than conceptual understanding do not accurately measure deeper learning outcomes (Kaur et al., 2016). When such items dominate an assessment, they can result in distorted evaluations of student results, essentially reducing the reliability and validity of the test.

Many MCTs require students to select the best answer among several possible options. While this format is intended to promote critical thinking, it often leads to confusion when more than one answer seems correct. Students with advanced knowledge may see nuances or exceptions not anticipated by the test designer and feel penalized for their deeper reasoning (Scouller, 1998). This issue uncovers a disparity between rigid answer key expectations and the

flexible nature of real-world reasoning. When students can defend alternative responses with sound logic, marking only one correct answer can be unfair. Such scenarios ask for more adaptive assessment formats, especially when the learning outcomes include analysis, interpretation, or judgment.

### ***Constructive alignment and MCTs***

Constructive alignment, a concept introduced by Biggs (1996), is a foundational principle in competence and outcome-based education. It promotes the alignment of learning outcomes, learning activities (tasks), and assessment methods, ensuring that all elements of the curriculum work coherently toward the development of intended student competencies. Within this approach, the use of multiple-choice questions represents a complex and often controversial aspect, criticized for potentially undermining deep learning.

Learning outcomes describe what students should know, understand, and be able to do after a learning experience (Schmidt, 2019). In well-conceived study programs, assessment tasks should provide valid evidence of the achievement of these outcomes. MCTs can be a suitable assessment method when learning outcomes target lower to mid-level cognitive skills, such as remembering, understanding, and applying, according to Bloom's revised taxonomy (Krathwohl, 2002). However, for outcomes emphasizing critical thinking, problem-solving, or synthesis, MCTs may fall short unless they are carefully constructed to test application and analysis, often through scenarios, data interpretation, or sequential reasoning (Pepple et al., 2010). This demands not only pedagogical expertise but also time and resources to develop high-quality items that truly reflect the complexity of intended learning outcomes.

Learning tasks are designed to help students achieve the intended outcomes. If such tasks promote only the passive absorption of information, such as lecture memorization or rote learning, the use of multiple-choice tests, which primarily assess factual recall, may appear appropriately aligned; however, this alignment can be misleading, as it reinforces surface-level learning. Passive intake of information refers to a learning approach in which students read material without engaging in meaningful interaction, analysis, or reflection. This might include watching lectures or reading texts without questioning, discussing, or applying the content. Specified learning environments tend to foster short-term memorization rather than deep understanding.

In a constructively aligned course, learning tasks should challenge students to explore, reflect, and apply knowledge, activities that MCTs often

do not mirror. To preserve alignment, academics need to ensure that learning tasks prepare students for the types of cognitive engagement required in the MCTs (Jovanovska, 2018). For example, if MCTs are designed to assess problem-solving, learning tasks should include guided problem-solving sessions, case analysis, or simulated labs. Moreover, the role of formative MCTs as learning activities should not be underestimated. When used during the learning process, MCTs can provide timely feedback and stimulate reflection. Online quiz tools, for example, allow students to engage in self-assessment and identify gaps in their understanding, a valuable practice in promoting self-regulated learning.

Assessment tasks must provide accurate evidence of students' achievement of learning outcomes. MCTs, valued for their objectivity and scalability, especially in large cohorts, are particularly effective in diagnosing misconceptions, offering feedback, and standardizing evaluation across students and teachers (Butler et al., 2007). Yet, the reliance on MCTs can lead to disillusionment when they are misused: for example, when they are the only form of assessment in courses where outcomes require higher-order thinking. In such cases, there is a misalignment; while the course may claim to foster critical thinking, the assessment tests just memory or recognition, thereby sending confusing indications to students about what is truly valued.

The objectivity of MCTs, often seen as strength, becomes a limitation when depth and nuance are required. Students may learn to "game the test" rather than engage meaningfully with the content. Additionally, overuse of MCTs can reduce opportunities for students to practice constructing arguments, applying theory to new contexts, or engaging in collaborative inquiry.

## Results

### *Practices from Romania and Moldova*

The educational assessment landscape in Romania and Moldova reflects both similarities and differences. In Romania, multiple-choice testing gained prominence particularly after the country's integration into European educational frameworks following EU accession in 2007. The National Assessment and the Baccalaureate examinations increasingly incorporated multiple-choice questions as part of efforts to standardize assessment procedures. This shift was motivated by practical considerations including the need for efficient evaluation of large student populations and the desire to reduce subjective grading variations across different regions of the country.

Multiple-choice testing was adopted in Moldova, primarily through educational reforms beginning in the early 2000s in secondary education and continuing in 2010 with introduction of competence-based curricula in Higher Education, often influenced by programs' designers, seeking to modernize the country's assessment practices. The assessment system starts a transition from the focus on oral examinations, written essays, and practical laboratory work to the emphasis on tests featuring multiple-choice items, short-answer questions, matching exercises, and fill-in-the-blank tasks.

Both countries face similar challenges regarding multiple-choice assessment validity. Research results from Romanian teachers indicate persistent concerns about whether standardized tests adequately measure student competence, particularly in fields requiring complex reasoning and practical skills. Romanian academics have expressed skepticism about over-reliance on multiple-choice formats, especially in humanities disciplines where nuanced interpretation and argument construction are valued.

In Moldova, teachers often combine multiple-choice questions with short-answer and essay components in an attempt to balance efficiency with assessment depth. This hybrid approach reflects ongoing efforts to preserve valuable aspects of their educational heritage while adopting standardized assessment methods. Assessment practices in higher education institutions exhibit significant variation, both across universities and within individual faculties. In many cases, standardized tests, administered either manually or in digital formats, continue to dominate as the primary method of evaluation. Although some faculties maintain the use of alternative assessment strategies, such as written assignments or project-based evaluations, the use of oral examinations has declined considerably compared to two decades ago. This shift reflects broader trends toward standardization, often at the expense of more personalized, discursive forms of assessment.

Romanian university teachers have developed a relatively high level of competence in designing and implementing various forms of tests, particularly multiple-choice and structured-response formats. Increasingly, academics are promoting the use of digital testing platforms, enhancing efficiency and accessibility in the assessment process. In some cases, artificial intelligence tools are also being employed to support the development, editing, and analysis of test items, reflecting a growing interest in integrating technology into educational assessment (Grosbeck et al., 2024).

In contrast, university teachers from Moldova have less experience in designing well-formulated test questions and appropriate answer options. Although many teachers have participated in training sessions through various educational projects, related to digital testing platforms, the practical

implementation of digital assessments remains limited across numerous institutions. As a result, the overall quality of test design often lacks the necessary rigor in terms of validity, clarity, and alignment with learning outcomes. This gap highlights the need for continued professional development focused specifically on assessment design and digital evaluation practices. Moreover, the use of artificial intelligence in test design and administration is still in its early stages within universities from Moldova. Only a small number of enthusiastic and tech-savvy educators have begun experimenting with AI tools to generate or refine test items. These efforts, while promising, remain isolated and are not yet integrated into institutional practices or supported through broader strategic initiatives. As such, the potential of AI to enhance assessment quality and efficiency is largely unexplored in the country higher education context.

Students' preparedness for taking tests is generally not a major concern, as they tend to adapt quickly to different assessment formats. What proves to be more critical is the clarity of test descriptions, the availability of detailed guidelines for answering, and the quality of feedback provided after the assessment. In Romanian universities, these elements are typically well-articulated, contributing to a more transparent and supportive testing environment. In contrast, Moldovan teachers often do not provide structured feedback or clear instructions, which can hinder students' ability to fully understand and engage with the test content. Additionally, biases in question formulation and weak correlation between questions and answer options can lead to confusion and misinterpretation, further compromising the fairness and effectiveness of the assessment.

### ***Qualitative findings from focus group data***

The presented findings of the study are related directly to the research purpose: to evaluate the extent to which multiple-choice tests reflect objective and pedagogically sound assessment practices in higher education based on literature research and real-life testing practices from Romania and Moldova. The results are linked to four key themes that emerged from both the literature analysis and the focus group discussions: (1) perceived objectivity and reliability, (2) cognitive depth and learning outcomes, (3) fairness and bias in test design, and (4) stakeholder preferences for assessment methods.

### **Perceived objectivity and reliability**

The literature widely supports the claim that MCTs offer standardized scoring, which is often equated with objectivity (Simkin et al., 2005). Teachers consider multiple-choice tests as objective tools for assessing student

knowledge because they are scored automatically and consistently: answers are either right or wrong, eliminating subjective judgment in grading. This objectivity is seen as a major advantage, as it helps standardize assessment across large groups and reduces the risk of bias that can occur in open-ended or essay-based evaluations.

This perspective was echoed by faculty members in the focus group, who emphasized that automated scoring eliminates potential grader bias and ensures consistency across large cohorts. However, many teachers recognize limitations: while the grading process is considered objective, the questions themselves may not always accurately measure deeper understanding or critical thinking skills, especially if poorly constructed (Haladyna & Rordiguez, 2013). Some academics worry that multiple-choice tests reward guessing and may not reveal true mastery of complex material, which can affect the perceived validity of the results. Despite these concerns, most teachers acknowledge that well-designed multiple-choice tests can provide reliable, efficient, and useful information about student knowledge when used appropriately.

### **Cognitive depth and learning outcomes**

MCTs were often criticized, both in the literature and by focus group participants, for promoting surface-level learning, particularly factual recall and recognition (Biggs & Tang, 2011). Both students and teachers agreed that MCTs rarely assess higher-order thinking skills such as analysis, synthesis, or critical reasoning.

Many students remarked: *“You can pass some of these tests by guessing or pattern recognition, not because you really understand the material. After a short time, you cannot remember what you learn”*.

The teachers acknowledged the practicality of MCTs for large groups but expressed concern that they fail to capture nuanced student understanding.

### **Fairness and bias in test design**

A significant point of convergence between literature and participant feedback concerned the design flaws and linguistic biases that can compromise test fairness (Solano-Flores & Nelson-Barber, 2001). Several students reported confusion caused by vague formulation or technical jargon that assumed background knowledge not evenly distributed across the class. Additionally, students raised concerns about the “best answer” format, especially when more than one option appeared defensible.

*“Sometimes I can argue for two answers depending on interpretation. Choosing just one feels arbitrary,”* said many students.

This finding aligns with previous research emphasizing that MCTs can unfairly penalize students who apply deeper reasoning than expected by the test designer (Scouller, 1998).

Some students from Moldova pointed out that bias in multiple-choice tests may arise from the subjectivity of the test authors. The subjectivity is hidden in:

- irrelevant formulations or questions not aligned with learning outcomes;
- poorly written or ambiguous distractors (incorrect answer choices) that mislead even well-prepared students, introducing unfairness into the assessment;
- variation in cognitive load or complexity of questions, yet each item is often assigned the same number of points, leading to an imbalance in the overall scoring;
- unequal weighting of content, paired with uniformly distributed scores, that results in misrepresentation of a student's actual understanding.

This evidence factors highlight the demand for rigorous test design and validation procedures, as unaddressed subjectivity can undermine the reliability, fairness, and interpretive accuracy of multiple-choice assessments.

### **Stakeholder preferences for assessment methods**

Both focus groups expressed the necessity for more balanced assessment practices, incorporating essays, oral presentations, and project-based evaluations, alongside MCTs. While teachers appreciated the administrative efficiency of MCTs, particularly in large-enrollment courses, they also acknowledged that alternative formats offer a more holistic proof of student learning.

*“Multiple-choice tests have their place, but it should be part of an assessment strategy with diversified methods,”* some teachers concluded.

Likewise, students advocated for a combination of methods that would allow them to demonstrate understanding in more expressive and analytical ways, reducing dependence on test-taking strategies. Almost all students expressed dissatisfaction with the test-based assessment, mostly due to the excessive workload and time pressure. They reported having to answer a large number of questions in a very limited time, which prevented them from reading the items carefully or reflecting on their answers. Many noted that

they were unable to revisit previous questions, further increasing their stress. By the end of the examination, students felt mentally fatigued and unable to maintain focus, which negatively affected their results and accuracy.

Students' approaches to learning can be categorized into two distinct types. The first category includes surface learners called also *surfer learners*, who are generally satisfied with test-based assessments. They tend to focus on rote memorization, often preparing by accessing previous cohorts' tests and memorizing the correct answers without engaging deeply with the content. In contrast, *deep learners* aim to understand underlying concepts, question assumptions, and construct meaning through critical thinking. For these students, the limitations of multiple-choice tests, especially when poorly designed, can hinder their ability to demonstrate understanding and may even demotivate them. Paradoxically, surface learners often achieve higher scores than deep learners in such formats, due to the assessment structure rewarding recognition over reasoning.

The key findings from the qualitative analysis are presented in the table 1.

*Table 1. Emergent themes from the qualitative analysis.*

<b>Theme</b>	<b>Strengths of MCTs</b>	<b>Limitations of MCTs</b>	<b>Stakeholders' views</b>
<b>Objectivity &amp; reliability</b>	Standardized, unbiased scoring	Perceived objectivity may be misleading	Valued for efficiency
<b>Cognitive depth</b>	Efficient for large groups	Promote surface learning	Insufficient for deep learning
<b>Fairness &amp; bias</b>	Consistent grading	Vulnerable to design flaws and bias	Concerns about unfairness
<b>Assessment preferences</b>	Practical for administration	Stressful, time-pressured, limited feedback	Need for mixed methods

*Source: authors' own concept*

These findings suggest that while MCTs are valued for their efficiency and reliability, their perceived objectivity may be misleading when test items are poorly designed or when they fail to align with complex learning outcomes. The reliance on recognition over reasoning, the vulnerability to guessing strategies, and the potential for linguistic and cultural bias all point to the limitations of MCTs as comprehensive assessment tools.

## Discussions and limitations of research

The findings from both the literature and focus group discussions confirm that multiple-choice tests are widely regarded as objective, mostly due to their standardized scoring and efficiency in large-scale assessments. This objectivity is valued in higher education contexts of Romania and Moldova, where current institutional tendencies, resource constraints and large student cohorts make automated, impartial grading highly attractive. Academic staff recognizes that MCTs reduce the risk of subjective bias and grading disputes, contributing to a perception of fairness and transparency in assessment. Nevertheless, this objectivity is not without its limitations. While MCTs are good in assessing factual recall and lower-order cognitive skills, their capacity to measure deeper understanding, application, and critical thinking remains questionable. Both students and teachers in the focus groups expressed concerns that MCTs often fail to capture the full spectrum of learning outcomes, especially those associated with higher-order thinking. This aligns with the literature, which cautions that an overreliance on MCTs can create a misleading sense of comprehensive evaluation. Students from Moldova express the opinion that bias in the MCTs may be caused by the subjectivity of the test authors.

Given the objectives of this article, the decision to rely solely on qualitative methods is both intentional and methodologically appropriate. The central aim — to explore whether multiple-choice tests represent true objectivity or a form of educational delusion — require a deep examination of how assessment items are designed, interpreted, and aligned with learning outcomes. Quantitative approaches, while useful for measuring test results, cannot reveal underlying biases, irrelevant content, misleading distractors, or misalignment with learning outcomes. In contrast, qualitative methods allow for a context analysis across different universities and disciplines, capturing the diversity of assessment practices and their pedagogical rationales. Thus, the use of qualitative inquiry not only uncovers what quantitative data might obscure but also supports a more nuanced and reflective evaluation of assessment in higher education.

Quantitative analysis alone cannot reveal whether assessment tasks are truly aligned with the intended learning outcomes. The statistical measures item difficulty or test reliability, but they do not assess the relevance or cognitive depth of the test items in relation to what students are expected to learn. As Biggs & Tang (2011) emphasize, alignment between assessment and learning outcomes is essential for effective teaching and valid evaluation, and this can only be judged through qualitative review of the content and purpose

of each task. Without such alignment, even a statistically reliable test may fail to measure meaningful student achievement.

Quantitative methods, such as calculating test scores, item difficulty, or discrimination indices, provide numerical insights into performance but cannot reveal potential biases embedded in the questions or answer options. These methods often overlook subtle issues such as bias, ambiguous wording, or misleading distractors, which may disadvantage certain groups of students. Statistical analysis alone cannot account for the contextual and interpretive aspects of assessment that influence how students understand and respond to test items. Therefore, a qualitative review of multiple-choice questions is essential to uncover flaws in content relevance, language clarity, or fairness, which purely numerical data may obscure (Rashwan et al., 2024).

The qualitative analysis within this research is focused on four central subjects: 1) perceived objectivity and reliability; 2) cognitive depth and learning outcomes; 3) fairness and bias in tests design; 4) stakeholder preferences for assessment methods. While this comparative analysis offers valuable insights into assessment practices in Romania and Moldova, it is essential to acknowledge several methodological and contextual limitations that may affect the scope and wider relevance of the results.

### **Sample size and representativeness**

The empirical component of this research is based on two focus groups: one with university students ( $n = 60$ ) and one with academic staff ( $n = 30$ ) from several selected universities in Romania and Moldova. The sample ensured diversity in academic backgrounds and experience with MCTs, but it is relatively small and context-specific that limits the generalization of findings to all higher education institutions in these countries or beyond. Also, institutional policies, assessment cultures, and educational traditions may differ significantly across regions that restrict the applicability of conclusions to other international contexts.

### **Complexity of assessment practices**

Assessment strategies in higher education are multifaceted and influenced by numerous factors, including institutional policy, resource availability, and course specificity. The qualitative analysis would not capture the full complexity of variables, and the diverse assessment practices across all institutions. Moreover, a significant limitation is the lack of open access to institutional information regarding assessment practices and specific assessment tasks, which constrains a more in-depth and comparative analysis.

### **Focus on higher education**

The analysis is based on the context of higher education, reflecting the specific focus and interest of the authors. Even the study does not cover secondary or vocational education preliminary observations suggest that similar challenges and perceptions regarding multiple-choice assessment exist across these educational levels in both Romania and Moldova. Future research could usefully extend this investigation to secondary and vocational settings, allowing for a more comprehensive understanding of assessment practices across the broader educational spectrum.

### **Evolving educational landscape**

The rapid evolution of digital assessment tools, online learning environments, and educational technologies may alter the role and perception of MCTs in the near future. The discussion is based on current practices and may require updating as new assessment paradigms emerge. Further research is needed to evaluate the effectiveness of other assessment formats in relation to MCTs.

There is a growing awareness among both teachers and students of the need to reimagine assessment as a multi-dimensional process that blends testing with more authentic, qualitative evaluations. For higher education institutions aiming to foster critical thinking and deep learning, a more integrated assessment framework is recommended.

### **Implications for the policy**

In the absence of national strategies for e-assessment and e-learning in Higher Education of both Romania and Moldova, it is essential to develop coherent policy frameworks that address these gaps. Such strategies should adopt an integrated approach that aligns pedagogical principles with the effective use of digital tools for teaching, learning and assessment. These documents must also prioritize systematic teacher training to ensure academic staff are equipped with the necessary pedagogical and digital competencies. Furthermore, the strategies should incorporate mechanisms for quality assurance, particularly regarding the design and implementation of assessment tasks, and ensure alignment at both the program and course levels within the curriculum. These coordinated efforts are critical to enhancing the effectiveness, fairness, and relevance of digital assessment in higher education.

## Conclusions

This article set out to examine the objectivity and pedagogical value of multiple-choice tests in higher education. An extensive literature review was combined with qualitative focus group discussions involving both students and academic staff from Romania and Moldova that resulted in a nuanced understanding of the strengths and limitations of MCTs as assessment methods.

The findings confirm that MCTs are widely perceived as objective and efficient, especially in large-scale educational settings where standardized scoring and automation are highly valued. MCTs assess lower-order cognitive skills, such as factual recall and basic understanding that makes them a practical choice for many institutions. The research also highlights significant limitations that challenge the use of MCTs as comprehensive tool of student learning. The so-called objectivity of MCTs is largely limited to the scoring process, while their capacity to assess higher-order thinking, deep understanding, and applied skills remains questionable. Both literature and focus group participants noted that MCTs often encourage surface learning and test-taking strategies, such as guessing or pattern recognition, rather than deep mastery of the subject matter. Poorly designed questions further undermine the validity and reliability of these assessments, sometimes penalizing students with advanced cognitive skills or creating confusion through ambiguous wording.

The research underscores the risk of misalignment between MCTs and intended learning outcomes, particularly when those outcomes emphasize critical thinking, creativity, or problem-solving. While constructive alignment theory supports the use of MCTs for certain types of knowledge, it also cautions against their overuse in contexts where deeper learning is the goal.

Overall, multiple-choice tests offer benefits in terms of objectivity, scalability and time saving for grading, but they should not be considered the sole or primary method of assessment in higher education. To ensure a valid and holistic evaluation of student achievement, MCTs must be carefully designed and complemented by alternative assessment formats, such as essays, case studies, portfolios, or oral examinations; that are appropriate for the demonstration of higher-order skills and deeper understanding.

For policy and practice, the findings suggest the need for ongoing professional development in assessment design, greater attention to the alignment of assessment methods with learning outcomes, and a balanced approach that integrates both objective and subjective assessment methods.

Future research should expand the empirical base through larger, more diverse samples and explore innovative assessment strategies that better

capture the complexity of student learning in contemporary higher education. Multiple-choice tests, being valuable for their efficiency, create an illusion of comprehensive assessment. Their effective use depends on a balanced, constructively aligned assessment approach that develops the range of competencies expected in higher education. Thus, although MCTs remain a popular choice for large-scale assessments, their continued popularity may reflect institutional convenience more than pedagogical alignment. The results show that relying solely on MCTs risks narrow the scope of learning assessment and undermines efforts to foster higher-order thinking and authentic knowledge construction.

### **Recommendations**

Based on a review of the literature and insights gathered through focus group discussions, the authors present the following recommendations aimed at improving assessment practices and addressing the limitations associated with multiple-choice tests (MCTs) in higher education institutions across Romania, Moldova, and comparable contexts. These recommendations are mainly directed toward university policymakers, curriculum designers, academic staff, and institutional quality assurance bodies responsible for curriculum development and assessment strategy implementation.

### **Multiple-choice tests as part of a balanced assessment strategy**

University academics and curriculum designers should ensure that MCTs are integrated thoughtfully within a diversified assessment framework. Assessments must be closely aligned with intended learning outcomes, especially when targeting complex competencies such as critical thinking, creativity, and problem-solving. Combining MCTs with other formats, including essays, case studies, portfolios, and oral examinations, allows for a more comprehensive assessment of students' mastery across both foundational knowledge and higher-order cognitive skills.

### **Constructive alignment in curriculum design**

Curriculum committees and quality assurance bodies are encouraged to continuously evaluate the coherence between learning outcomes, learning methods, and assessment approaches. This constructive alignment ensures that teaching and assessment mutually reinforce meaningful and deep learning. Active learning tasks engage students in analysis and knowledge application and prepare them for diverse assessment types beyond standard MCTs.

### **The quality of multiple-choice test design**

Academic departments and decision makers should invest in ongoing professional development focused on the design of high-quality MCTs. Training should emphasize constructing items that reflect cognitive complexity, ensure cultural fairness, and maintain alignment with learning outcomes. Systematic tests analysis should be established as a core quality assurance mechanism to identify ambiguous, biased, or ineffective questions, with mechanisms for revising or removing such items. Additionally, peer review processes for test questions should be established to promote rigor and diverse perspectives in assessment design.

### **Risks of superficial learning and guessing**

Teachers are advised to design MCTs that go beyond mere recall by using scenario-based, application, and analysis questions. This approach reduces the likelihood of guessing and encourages deep understanding. Distractors should be thoughtfully elaborated to require students' meaningful engagement with the content, avoiding easily exploitable patterns that test-wise students might use to their advantage.

### **Transparency and fairness in assessment**

Teachers and test developers must communicate clearly with students regarding the purpose, format, and scoring criteria of MCTs, explaining them the assessment strategy. Teachers should provide helpful and timely feedback to students that highlight what they did well and where they can improve; this is important for supporting students' learning and motivation, beyond just their grades.

### **Ongoing research and institutional dialogue**

Higher education institutions and policymakers should support empirical research on assessment practices, ideally involving larger and more diverse student samples. Such research will validate current findings and foster innovation in assessment methods. Furthermore, creating collaborative forums for sharing best practices, challenges, and policy developments across universities in Romania, Moldova, and the surrounding region can enhance collective capacity for assessment quality assurance.

These recommendations aim to guide stakeholders across multiple levels of higher education institutions in enhancing the effectiveness, fairness, and educational impact of assessment practices, with particular attention to

the responsible use of multiple-choice tests within a broader evaluation framework.

This work was funded by UEFISCDI PN-IV-P8-8.3-ROMD-2023-0149: Perspective ale evaluării digitale în învățământul superior din România și Moldova: o analiză comparativă (Perspectives on Digital Assessment in Higher Education in Romania and Moldova: A Cross-Country Analysis), <http://deva.uvt.ro>

---

## References

---

- Alordiah, C., & Oji, J. (2024). Test equating in educational assessment: A comprehensive framework for promoting fairness, validity, and cross-cultural equity. *Asian Journal of Assessment in Teaching and Learning*, 14(1). <https://doi.org/10.37134/ajatel.vol14.1.7.2024>
- Bateson, A., & Dardick, W. R. (2020). A comparison of the two-option versus the four-option multiple-choice item: A case for fewer distractors. *Personnel Assessment and Decisions*, 6(3), Article 5. <https://doi.org/10.25035/pad.2020.03.005>
- Berg, P., & Singh, A. (2024). Myths of official measurement: Limits to test-based education reforms with weak governance. *Journal of Public Economics*, 239. <https://doi.org/10.1016/j.jpubeco.2024.105246>
- Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education*, 32, 347–364. <https://doi.org/10.1007/BF00138871>
- Biggs, J., & Tang, C. (2011). *Teaching for quality learning at university* (4<sup>th</sup> ed.). Open University Press.
- Butler, A. C. (2018). Multiple-choice testing in education: Are the best practices for assessment also good for learning? *Journal of Applied Research in Memory and Cognition*, 7(3), 323–331. <https://doi.org/10.1016/j.jarmac.2018.07.002>
- Butler, A. C., Karpicke, J. D., & Roediger, H. L. III. (2007). The effect of type and timing of feedback on learning from multiple-choice tests. *Journal of Experimental Psychology: Applied*, 13(4), 273–281. <https://doi.org/10.1037/1076-898X.13.4.273>
- Crețu, D. M., & Grosseck, G. (2025). A bibliometric analysis of Romanian educational research in Web of Science: Trends, challenges, and opportunities for global integration. *Education Sciences*, 15(3), 358. <https://doi.org/10.3390/educsci15030358>
- DiBattista, D., & Kurzawa, L. (2011). Examination of the quality of multiple-choice items on classroom tests. *The Canadian Journal for the Scholarship of Teaching and Learning*, 2(2). <https://doi.org/10.5206/cjsotl-rcacea.2011.2.4>

- Douglas, K. A., & Purzer, Ș. (2015). Validity: Meaning and relevancy in assessment for engineering education research. *Journal of Engineering Education*, 104(2), 108-118. <https://doi.org/10.1002/jee.20070>
- Dumbraveanu, R. (2006). *Evaluarea studenților: Strategii și metode*. Universitatea Pedagogică de Stat “Ion Creangă”.
- Dumbraveanu, R., & Peca, L. (2022). E-learning strategy in the elaboration of courses. *International Conference on Virtual Learning*, 17, 15–26. <https://doi.org/10.58503/icvl-v17y202201>
- EHEA. (2020). *Bologna Process: Bologna beyond 2020 – Fundamental values of the EHEA*. Rome Ministerial Conference, European Higher Education Area. [https://eha.info/Upload/Rome\\_Ministerial\\_Communique.pdf](https://eha.info/Upload/Rome_Ministerial_Communique.pdf)
- ENQA. (2015). *Standards and guidelines for quality assurance in the European Higher Education Area* (3rd ed.). European Association for Quality Assurance in Higher Education. [https://www.enqa.eu/wp-content/uploads/2015/11/ESG\\_2015.pdf](https://www.enqa.eu/wp-content/uploads/2015/11/ESG_2015.pdf)
- European Commission. (2023). *Council Recommendation on improving the provision of digital skills*. <https://education.ec.europa.eu/focus-topics/digital-education/action-plan/council-recommendation-improving-the-provision-of-digital-skills>
- Goss, H. (2022). Student learning outcomes assessment in higher education and in academic libraries: A review of the literature. *The Journal of Academic Librarianship*, 48(2). <https://doi.org/10.1016/j.acalib.2021.102485>
- Grossec, G., Bran, R. A., & Țîru, L. G. (2024). Digital assessment: A survey of Romanian higher education teachers’ practices and needs. *Education Sciences*, 14(1), 32. <https://doi.org/10.3390/educsci14010032>
- Haladyna, T. M., & Rodriguez, M. C. (2013). *Developing and validating test items* (1<sup>st</sup> ed.). Routledge. <https://doi.org/10.4324/9780203850381>
- Jovanovska, J. (2018). Designing effective multiple-choice questions for assessing learning outcomes. *Infotheca*, 18(1), 25–42. <https://doi.org/10.18485/infotheca.2018.18.1.2>
- Kaur, M., Singla, S., & Mahajan, R. (2016). Item analysis of in-use multiple choice questions in pharmacology. *International Journal of Applied & Basic Medical Research*, 6(3), 170–173. <https://doi.org/10.4103/2229-516x.186965>
- Kent-Waters, J, Seago, O., Smith, L. (2018). *The compendium of assessment techniques in higher education: From students' perspectives*. Leeds Institute for Teaching Excellence. <https://teachingexcellence.leeds.ac.uk/wp-content/uploads/sites/89/2018/10/PUGHcompendiumcomplete.pdf>
- Krathwohl, D. R. (2002). A revision of Bloom’s taxonomy: An overview. *Theory Into Practice*, 41(4), 212–218. <http://www.jstor.org/stable/1477405>

- Kuechler, W. L., & Simkin, M. G. (2010). Why is performance on multiple-choice tests and constructed-response tests not more closely related? Theory and an empirical test. *Journal of Innovative Education*, 8(1), 55–73. <https://doi.org/10.1111/j.1540-4609.2009.00243.x>
- McKenna, P. (2018). Multiple choice question: Answering correctly or knowing the answers. *International Conference e-Learning 2018*. <https://files.eric.ed.gov/fulltext/ED590289.pdf>
- Ministerul Educației și Cercetării. (2015). *Strategia națională pentru învățământ terțiar 2015–2020*. [https://www.edu.ro/sites/default/files/fisiere%20articole/Strategie\\_invt\\_ertiar\\_2015\\_2020.pdf](https://www.edu.ro/sites/default/files/fisiere%20articole/Strategie_invt_ertiar_2015_2020.pdf)
- Nicol, D. (2007). E-assessment by design: Using multiple-choice tests to good effect. *Journal of Further and Higher Education*, 31(1), 53–64. [https://www.liverpool.ac.uk/media/livacuk/cll/eddev-files/iteach/pdf/nicole\\_assessment\\_by\\_design\\_using\\_MCOs.pdf](https://www.liverpool.ac.uk/media/livacuk/cll/eddev-files/iteach/pdf/nicole_assessment_by_design_using_MCOs.pdf)
- Nitko, A. J., & Brookhart, S. M. (2014). *Educational assessment of students*. (8<sup>th</sup> ed.) Pearson Education
- Oc, Y., & Hassen, H. (2025). Comparing the effectiveness of multiple-answer and single-answer multiple-choice questions in assessing student learning. *Marketing Education Review*, 35(1), 44–57. <https://doi.org/10.1080/10528008.2024.2417106>
- Peple, D. J., Young, L. E., & Carroll, R. G. (2010). A comparison of student performance in multiple-choice and long essay questions in the MBBS stage I physiology examination at the University of the West Indies. *Advances in Physiology Education*, 34(2), 86–89. <https://doi.org/10.1152/advan.00087.2009>
- Pitt, E., & Quinlan, K. (2022). *Impacts of higher education assessment and feedback policy and practice on students: A review of the literature 2016–2021*. University of Kent. <https://www.advance-he.ac.uk/knowledge-hub/>
- Popham, W. J. (2008). *Transformative assessment*. ASCD. <https://www.perlego.com/book/3292589/transformative-assessment-pdf>
- Rashwan, N. I., Aref, S. R., Nayel, O. A., & Rizk, M. H. (2024). Post examination item analysis of undergraduate pediatric multiple-choice questions: Implications for developing a validated question bank. *BMC Medical Education*, 24, 168. <https://doi.org/10.1186/s12909-024-05153-3>
- Republic of Moldova, Government. (2023). *Strategia de dezvoltare „Educația 2030”* (Government Decision No. 114/2023). Ministry of Education and Research. [https://www.legis.md/cautare/getResults?doc\\_id=136600&lang=ro](https://www.legis.md/cautare/getResults?doc_id=136600&lang=ro)
- Rios, J. A., Deng, J., & Ihlenfeldt, S. D. (2022). To what degree does rapid guessing distort aggregated test scores? A meta-analytic investigation. *Educational Assessment*. <https://doi.org/10.1080/10627197.2022.2110465>

- Rodriguez, M. C. (2003). Construct equivalence of multiple-choice and constructed-response items: A random effects synthesis of correlations. *Journal of Educational Measurement*, 40(2), 163–184.  
<https://www.jstor.org/stable/1435344>
- Rodriguez-Torrealba, R., Garcia-Lopez, E., & Garcia-Cabot, A. (2025). Joint generation of distractors for multiple-choice questions: A text-to-text approach. *Computers, Materials & Continua*, 83(2), 1683–1705.  
<https://doi.org/10.32604/cmc.2025.062004>
- Schmidt, H. U. (2019). Learning outcomes: Core issues in higher education. In D. Staub (Ed.), *Quality assurance and accreditation in foreign language education* (pp. 233–247). Springer. [https://doi.org/10.1007/978-3-030-21421-0\\_12](https://doi.org/10.1007/978-3-030-21421-0_12)
- Scouller, K. (1998). The influence of assessment method on students' learning approaches: Multiple choice question examination versus assignment essay. *Higher Education*, 35(4), 453–472.  
<https://doi.org/10.1023/A:1003196224280>
- Simkin, M. G., & Kuechler, W. L. (2005). Multiple-choice tests and student understanding: What is the connection? *Journal of Innovative Education*, 3(1), 73–98. <https://doi.org/10.1111/j.1540-4609.2005.00053.x>
- Slepkov, A. D., Van Bussel, M. L., Fitze, K. M., & Burr, W. S. (2021). A baseline for multiple-choice testing in the university classroom. *SAGE Open*, 11(2).  
<https://doi.org/10.1177/21582440211016838>
- Solano-Flores, G., & Nelson-Barber, S. (2001). On the cultural validity of science assessments. *Journal of Research in Science Teaching*, 38(5), 553–573.  
<https://doi.org/10.1002/tea.1018>
- Tarrant, M., & Ware, J. (2012). A comparison of the psychometric properties of three- and four-option multiple-choice questions in nursing assessments. *Nurse Education Today*, 32(4), 345–349.  
<https://doi.org/10.1016/j.nedt.2011.05.002>