INTEGRATED SUMMATIVE ASSESSMENT AS A TOOL FOR ENHANCING CRITICAL THINKING COMPETENCE IN SCHOOL EDUCATION

Tursunaliyev Shaxzod Sherali o‘g‘li
Termiz davlat pedagogika instituti
mustaqil izlanuvchisi

Abstract. *The article examines integrated summative assessment as a pedagogical mechanism for developing students' critical thinking competence in school education. The study argues that critical thinking cannot be strengthened by end-of-term testing built only on reproduction of facts, closed questions, and isolated subject content. A more productive model is an integrated summative assessment that combines selected-response items with analytical writing, evidence-based explanation, problem solving, source comparison, oral defence, and reflective judgement within a shared rubric. On that basis, the paper clarifies the structure of critical thinking competence, identifies assessment principles that make higher-order reasoning visible, and describes the didactic conditions required for reliable use in school practice. Particular attention is given to the alignment of learning outcomes, task design, scoring criteria, and feedback procedures. The article concludes that integrated summative assessment becomes effective when it measures not only the final answer, but also the quality of argument, interpretation, transfer, justification, and self-correction demonstrated by the learner.*

Keywords: *summative assessment, integrated assessment, critical thinking competence, school education, higher-order thinking, performance task, rubric, educational outcomes*

Introduction

In contemporary school education, the quality of learning is increasingly judged not by the amount of memorized information, but by the learner's ability to interpret information, evaluate evidence, compare alternatives, make reasoned decisions, and transfer knowledge to unfamiliar situations. This shift has made critical thinking one of the central outcomes of modern education. Brookhart explains that assessment of higher-order thinking should move beyond recall and should be deliberately designed around novel material, visible reasoning, and scoring procedures that capture the quality of thought rather than the quantity of reproduced facts [1]. A related theoretical position is advanced by Dwyer, Hogan, and Stewart, who define critical thinking as a metacognitive process involving purposeful and reflective judgement, supported by skills, dispositions, and contextual knowledge [2]. Halpern also emphasizes that critical thinking is teachable and that students

develop it more productively when they work with argument analysis, decision making, problem solving, and evidence-based explanation instead of passive repetition [3].

Despite this broad agreement, school assessment practice often remains conservative. End-of-unit and end-of-term tests are still dominated by tasks that privilege recognition, short recall, or single-correct-answer routines. Ennis noted long ago that critical thinking assessment is difficult, but possible, and that the choice of format must depend on the precise purpose of assessment and the kind of thinking to be elicited [4]. Later studies confirmed that many school tests still underrepresent higher-order reasoning. Orhan and Çeviker Ay, while developing a critical thinking test for high school students, operationalized the construct through inference, evaluation of arguments, deduction, recognition of assumptions, and interpretation, thus showing that critical thinking can be assessed through clearly defined subskills rather than vague general impressions [5]. In a different context, Shafiyeva argued that overreliance on standardized end-of-semester testing narrows teaching, equates learners with scores, and weakens opportunities for comparison, discussion, inquiry, and analytical judgement [6].

For this reason, the present article focuses on integrated summative assessment as a more adequate tool for enhancing critical thinking competence in school education. The term is used here to describe a summative assessment model that combines several mutually reinforcing task types within a unified design: objective questions for basic accuracy, constructed responses for explanation, performance tasks for application, and reflective components for metacognitive control. Such integration does not cancel the summative function of certification. Rather, it broadens the evidence base on which final judgements are made. The main purpose of the article is to show why this model is pedagogically stronger than fragmentary testing, what structural components it should include, and under what conditions it can genuinely support students' critical thinking competence.

Main part

The first methodological issue concerns the structure of critical thinking competence itself. In school education, this competence should not be reduced to general intelligence or rhetorical fluency. It includes a cluster of demonstrable actions: identifying a problem, separating fact from opinion, detecting assumptions, evaluating the reliability of a source, recognizing weak arguments, comparing explanations, producing justified conclusions, and revising one's answer when contradictory evidence appears. The assessment model developed by Orhan and Çeviker Ay is useful here because it translates the broad idea of critical thinking into measurable subskills that can be adapted to school subjects [5]. Dwyer, Hogan, and Stewart add an important dimension by linking skill performance with reflective judgement and disposition toward thinking, which means that assessment must capture not only whether the learner answered, but how the learner reasoned [2]. Therefore,

a school that seeks to enhance critical thinking through summative assessment must begin by articulating a competence model with explicit indicators.

The second issue concerns the weakness of traditional summative assessment. When end-of-term testing privileges mechanical recall, students quickly discover that reproduction is more profitable than reasoning. The assessment system then silently teaches dependence on templates, not intellectual independence. A concrete illustration appears in the study by Rianti, Aziz, and Aulia, who found that higher-order thinking represented only 4.6 percent of the questions in the English summative assessments they analyzed, while lower-order items dominated the test structure [7]. Shafiyeva similarly shows that standardized testing systems can produce a culture in which teachers prepare students for likely questions instead of organizing inquiry-rich learning [6]. In such conditions, critical thinking is declared in curriculum documents but marginalized in real classroom accountability. The problem, then, is not summative assessment as such, but a narrow design logic that reduces achievement to short, easily scored fragments.

Integrated summative assessment offers a stronger alternative because it unites several forms of evidence under one competence-oriented framework. Its central principle is constructive alignment: the assessed task must correspond to the intended learning outcome, and the scoring criteria must correspond to the cognitive action demanded by the task. If the target outcome is critical thinking, then the learner must be required to interpret, compare, justify, critique, or solve, not simply remember. Brookhart's work is especially important here because it insists that higher-order assessment should involve novel materials, situations, or source sets that force students to think in the moment [1]. Ennis likewise warns that prepackaged instruments alone are insufficient and that well-designed teacher-made tasks remain necessary for valid assessment [4]. This leads to an important practical conclusion: an integrated summative assessment should combine closed items only for foundational knowledge, while the core score should come from tasks that make argument quality visible.

Table 1. Core components of an integrated summative assessment for critical thinking competence

Assessment component	**Typical student action**	**Sample evidence**	**Critical-thinking indicator**
Source analysis task	Compare two or more texts, data sets, or viewpoints	Marked comparison, identification of bias, evidence selection	Analysis and evaluation

Written argument	Formulate a claim and justify it with evidence	Structured response with reasons, examples, and conclusion	Inference and justification
Problem-based performance task	Apply subject knowledge to a new situation	Decision, explanation of steps, alternative solution	Transfer and problem solving
Oral defence or questioning	Respond to counterarguments and clarify reasoning	Short oral explanation, defence of answer	Argument appraisal
Reflective self-assessment	Evaluate the strength and limitation of one's own response	Brief reflection linked to rubric criteria	Metacognitive regulation

The internal design of integrated summative assessment should therefore be multi-component, but not chaotic. It is not enough to place a project beside a test and call the result integration. Each component must perform a distinct diagnostic role. Selected-response items can verify conceptual precision and prerequisite knowledge. Short constructed responses can require explanation and interpretation. An extended task can demand comparison of sources, solution of a real or simulated problem, or defence of a position. An oral mini-defence or reflective commentary can reveal whether the student genuinely understands the reasoning behind the written answer. When these components are joined by a common analytic rubric, the final summative judgement becomes more valid because it is based on multiple traces of thinking rather than a single test score. In this sense, integration refers not merely to diversity of format, but to the coherence of evidence collection.

A decisive condition of effectiveness is the scoring model. If teachers assign marks only for factual correctness, then even a rich task can collapse into a low-level exercise. Brookhart directly points out that the scoring scheme must require higher-order thinking in order for the assessment itself to do so [1]. Thus, analytic rubrics in school education should allocate distinct criteria for interpretation of information, relevance of evidence, quality of reasoning, logical consistency, originality of approach, and capacity for revision. Such rubrics also improve transparency for students because they show that a strong answer is not just the right answer, but a justified answer. At the same time, teachers need moderation procedures, exemplars, and common judgement sessions so that performance-based summative assessment remains reliable. Without this step, integrated assessment risks becoming intellectually valuable but administratively unstable.

Another necessary condition is pedagogical continuity. Critical thinking cannot be measured summatively at the end if it has not been practiced formatively during the learning process. Mkomele's study of secondary education highlights both the potential of open-ended questioning, presentations, rubric-based examination, and practical tasks, and the barriers created by weak teacher preparation and insufficient resources [8]. This means that integrated summative assessment must be supported by day-to-day classroom routines in which students regularly analyse texts, evaluate statements, justify conclusions, and respond to feedback. Otherwise, the final task will function less as a fair measure of competence and more as a surprise obstacle. In pedagogical terms, the summative stage should culminate a sequence of rehearsed intellectual actions rather than introduce them for the first time.

In school practice, integrated summative assessment also has an interdisciplinary advantage. Critical thinking competence rarely develops within the boundaries of one subject alone. A history task may require interpretation of evidence, a science task may require hypothesis testing, and a language task may require argument evaluation and textual comparison. When schools coordinate criteria across subjects, students begin to recognize that critical thinking is not a decorative phrase specific to one classroom, but a transferable mode of academic work. Halpern's argument about transfer is especially relevant here: students learn critical thinking more effectively when they encounter the same reasoning demands across varied contexts [3]. Therefore, integrated summative assessment becomes particularly powerful when schools adopt shared descriptors for evidence use, argument quality, and reflective judgement while still preserving disciplinary specificity in content.

At the same time, this model should not be romanticized. Integrated summative assessment demands more preparation time, better task writing, teacher training in rubric use, and organizational support from school leadership. It is easier to score a conventional test than to evaluate a comparative essay, oral defence, and reflective note. Yet convenience cannot be the principal criterion in education policy. If the declared objective is to prepare students for complex reasoning, then the assessment system must reward complex reasoning. Shafiyeva's critique of score-driven standardization and Rianti's evidence of HOTS underrepresentation both point to the same conclusion: what schools repeatedly test is what students will habitually learn to value [6; 7]. Integrated summative assessment changes that signal.

Conclusion

Integrated summative assessment should be regarded not as an optional methodological decoration, but as a necessary instrument for aligning school assessment with the educational goal of critical thinking competence. Its pedagogical value lies in the fact that it expands the evidence base of summative judgement and enables teachers to evaluate

reasoning, interpretation, transfer, and justification together with subject knowledge. The article has shown that critical thinking becomes visible only when assessment tasks are designed around meaningful intellectual action, when rubrics reward argument quality rather than mechanical reproduction, and when final assessment continues the logic of classroom practice. For school education, the most productive model is a coherent combination of objective items, analytical writing, problem-based performance, oral explanation, and reflective self-assessment. Such a model helps shift the culture of achievement from memorization toward judgement. Consequently, integrated summative assessment can genuinely enhance critical thinking competence when it is grounded in explicit criteria, interdisciplinary coordination, teacher preparation, and a commitment to assessing how students think, not merely what they can repeat.

FOYDALANILGAN ADABIYOTLAR

1. Brookhart S. M. How to Assess Higher-Order Thinking Skills in Your Classroom. Alexandria, VA: ASCD, 2010. 176 p.

2. Dwyer C. P., Hogan M. J., Stewart I. An integrated critical thinking framework for the 21st century // Thinking Skills and Creativity. 2014. Vol. 12. P. 43–52. DOI: 10.1016/j.tsc.2013.12.004.

3. Halpern D. F. Thought and Knowledge: An Introduction to Critical Thinking. 5th ed. New York: Psychology Press, 2014.

4. Ennis R. H. Critical thinking assessment // Theory into Practice. 1993. Vol. 32, No. 3. P. 179–186. DOI: 10.1080/00405849309543594.

5. Orhan A., Çeviker Ay Ş. Developing the Critical Thinking Skill Test for High School Students: A Validity and Reliability Study // International Journal of Psychology and Educational Studies. 2022. Vol. 9, No. 1. P. 130–142. DOI: 10.52380/ijpes.2022.9.1.561.

6. Shafiyeva U. Assessing Students' Minds: Developing Critical Thinking or Fitting into Procrustean Bed // European Journal of Education. 2021. Vol. 4, No. 2. P. 79–92. DOI: 10.26417/452bxv17s.

7. Rianti R., Aziz Z. A., Aulia M. Incorporating Higher Order Thinking Skills into English Summative Assessments // English Review: Journal of English Education. 2024. Vol. 12, No. 1. P. 353–360. DOI: 10.25134/erjee.v12i1.9301.

8. Mkomele K. I. Teachers' Perception on Critical Thinking in Secondary School in Tanzania: A Case Study of Mjimwema Ward: dissertation for the degree of Master of Education in Administration, Planning and Policy Studies. The Open University of Tanzania, 2015. 86 p.