



Akademik Eğitim ve Sosyal Bilimler Dergisi (AJESS Journal)

From Measuring Knowledge to Assessing Competencies Conceptual Transformations in Contemporary Educational Evaluation

Khouloud Merakha, Badji Mokhtar University - Annaba, P.O. Box 12, 23000 Annaba (Algeria)
khouloud.merakha@univ-annaba.dz

Salima Bounaidja Rachedi, Badji Mokhtar University - Annaba, P.O. Box 12, Annaba 23000 (Algeria)
Salima.bounaidja-rachedi@univ-annaba.dz
Laboratory of Poetics and Discourse Analysis LPAD

Article History

Received: 01/08/2026; Accepted: 24/09/2026; Published: 05/10/2026

Abstract:

Educational evaluation has undergone a substantial conceptual transformation: from the quantification of discrete learning outcomes toward the interpretation of evidence about learners' ability to mobilize knowledge, skills, and other resources in context. This shift became especially visible with competency-based curricula and raises a central theoretical question: Is competency-based evaluation an extension of educational measurement, or does it represent a break from it?

Using a theoretical-analytical approach, this article traces the conceptual development of measurement, evaluation, knowledge, competency, and integration situations. The analysis identifies transformations in the object, tools, and functions of evaluation. It argues that competency-based evaluation neither eliminates measurement nor reduces to it: measurement produces quantitative evidence, while evaluation integrates evidence with criteria and professional judgment to support decisions about learning. The resulting model clarifies the role of complex tasks, formative use of evidence, and teacher assessment literacy.

Keywords: educational evaluation; assessment; measurement; knowledge; competencies; competency-based education

Introduction

The competency-based approach has reshaped the relationship among curriculum, teaching, learning, and evaluation. Evaluation is no longer adequately conceived as a final procedure that assigns a score to what the learner remembers; it is increasingly understood as a process for eliciting and interpreting evidence across the learning sequence. The conceptual shift is therefore not simply from 'less measurement' to 'more evaluation.' It concerns what counts as evidence, how evidence is interpreted, and what decisions it is intended to support. In this article, evaluation is used as the broader process of judging educational achievement and informing decisions, whereas measurement refers to the rule-governed assignment of quantitative values to attributes of performance. Assessment is used when referring specifically to the collection and interpretation of evidence about learner performance. This distinction is important because the three processes overlap without being synonymous. The central question is therefore: Is competency-based evaluation an extension of educational measurement, or does it involve a genuine epistemological break? The article addresses this question through conceptual analysis of the object, tools, and functions of evaluation.

The analysis proceeds through three linked movements: (1) from measurement to evaluation, (2) from knowledge to competencies, and (3) from isolated test items to integrated evaluation situations. It then offers a contemporary synthesis of how measurement and judgment can be combined within valid competency-based assessment.

Analytical Approach and Scope

Because the article is conceptual rather than empirical, its analytical procedure consists of comparing definitions and theoretical models according to three dimensions: the object of evaluation, the form of evidence used, and the function assigned to the resulting judgment. The selected literature represents successive conceptual moments in the development of measurement, competence, integration pedagogy, and contemporary classroom assessment. The purpose is not to establish a chronological history of every theory of evaluation, but to identify the conceptual shifts that explain why competency-based evaluation cannot be reduced to the measurement of recalled knowledge. The analysis therefore uses comparison and synthesis to examine continuity as well as change. This scope is important: the article does not claim that all traditional tests assess only memorization, nor that every complex task provides valid evidence of competence. Rather, it examines the conditions under which different forms of evidence support different educational claims and decisions.

First: Conceptual Grounding: From Measurement to Evaluation

Before examining theoretical transformations in educational evaluation, it is necessary to distinguish measurement, assessment, and evaluation. They are related, but they answer different questions: measurement asks how much, or at what level, an attribute is present; assessment asks what evidence of learning can be elicited and interpreted; evaluation adds criteria and judgment in order to support a decision. Confusing these levels risks treating a score as if it were a complete educational judgment.

Measurement: Stevens's classic formulation defines measurement, in its broadest sense, as the assignment of numerals to objects or events according to rules.¹ In educational contexts, measurement produces scores or quantitative indicators for specified attributes. Its importance lies in making performance describable and comparable under explicit rules; however, a numerical result does not by itself determine the meaning, quality, or educational consequence of that result.

Evaluation: Evaluation is broader than measurement because it combines evidence with criteria in order to make a judgment about the extent to which educational objectives or intended outcomes have been achieved.² It may incorporate quantitative scores, qualitative descriptions, observations, or performance evidence. In this sense, measurement can contribute to evaluation, but evaluation also requires interpretation and decision-making.

A further distinction concerns the status of the evidence produced at each level. A measurement result is meaningful only in relation to the attribute and rules through which it was generated. An assessment interpretation goes a step further by connecting observed performance to a claim about learning. Evaluation adds another layer by relating that interpreted evidence to educational purposes, standards, and consequences. This sequence shows why numerical precision cannot substitute for conceptual clarity. A highly reliable score may still provide weak evidence for a competency if the task samples only a narrow part of the intended construct. Conversely, rich classroom evidence may be educationally useful but insufficient for high-stakes certification if criteria are vague or scoring is inconsistent. The conceptual transformation of evaluation therefore involves not the abandonment of quantitative evidence, but a more explicit concern with the adequacy of the inference made from evidence to judgment. In competency-based education, this concern becomes central because the construct is broader than any single observable performance.

Educational evaluation has a set of functions summarized by De Landsheere³ as follows:

- Predictive function: Evaluation provides a forward-looking view of the learner. Does the learner possess what is needed to undertake a new subject? Can the learner succeed?
- Appraisal function: It monitors the learner's acquisitions and evaluates the extent of progress toward the stated objectives.
- Diagnostic function: It identifies the learner's strengths and weaknesses in order to choose the necessary study plan. Abdelaziz Quraish⁴ distinguished two basic functions, in addition to the traditional certification function:
- Guidance function (before the act of learning): It is concerned with identifying learners' readiness and qualifications.
- Adjustment function (during the act of learning): It is concerned with identifying shortcomings in order to address them.
- Certification function (after the act of learning): It is concerned with highlighting the learner's strengths.

¹ S. S. Stevens, "On the Theory of Scales of Measurement," *Science* 103, no. 2684 (1946), pp. 677-680, p. 677.

² Gilbert de Landsheere, *Dictionnaire de l'évaluation et de la recherche en éducation*, 2nd rev. and expanded ed., PUF, Paris, 1992, p. 111.

³ Gilbert de Landsheere, *Évaluation continue et examens: Précis de docimologie*, 5th rev. ed., Labor, Brussels, 1980, p. 15.

⁴ Abdelaziz Quraish, "Educational Evaluation: The Question of Nature and Function," *Education and Training Notebooks*, no. 4, February 2011, Morocco, p. 9.

This temporal sequence indicates that evaluation⁵ has the following types:

1. Diagnostic evaluation: It helps the teacher identify the level of readiness that learners possess and that enables them to reach a certain level of competency.
2. Formative evaluation: It is the evaluation of the educational process. It enables the teacher to identify strengths and shortcomings in the tools and strategies adopted in the teaching-learning process.
3. Certification evaluation: It is the stage in which a judgment is made about learners' learning and they are enabled to progress.

From the foregoing, the change is better described as an expansion in the nature and functions of evaluation rather than a total break with measurement. When measurement is restricted to quantitative description, evaluation goes beyond it by adding criteria, interpretation, and value judgment concerning the educational act. Its functions likewise extend beyond certification to guidance, diagnosis, and adjustment. Measurement therefore remains necessary in many evaluative contexts, but it is not sufficient for the broader decisions that evaluation requires.

This nested relationship also clarifies an important terminological difficulty. In practice, the same classroom activity may contain measurement, assessment, and evaluation at once. A teacher may assign a score to a task, interpret the errors and strategies displayed by the learner, and then decide whether additional support, progression, or certification is justified. The three operations can therefore occur in a single sequence while remaining conceptually distinct. Recognizing this distinction prevents two opposite errors: reducing evaluation to scoring, and treating professional judgment as independent of systematically collected evidence. A contemporary theory of educational evaluation must preserve both dimensions. It requires sufficiently structured evidence to support consistency, while retaining the interpretive dimension needed to judge learning in relation to educational aims and contextual demands.

Does this expansion at the level of function also affect the object and tools of evaluation? This is addressed in the following section.

From Knowledge to Competencies:

Traditional educational practice often privileged the measurement of discrete knowledge through tasks that rewarded recall and reproduction. In such contexts, the score functioned as the main basis for progression decisions. Competency-based approaches shift attention toward what learners can do with what they know. This does not make knowledge less important; rather, it changes its status from the endpoint of evaluation to a resource that must be selected, combined, and mobilized in context. The relationship between knowledge and competency therefore requires conceptual clarification.

Knowledge: It is objective knowledge, "information and content about the world that can be exchanged between individuals."⁶ It includes all the information received by the learner. Its evaluation does not go beyond determining the extent to which it can be recalled. This is consistent with the traditional concept of measurement.

Competency:

Before the concept of competency became established as an educational term, a decisive formulation appeared in linguistics through Chomsky's distinction between competence and performance. Linguistic competence referred to the speaker-hearer's underlying knowledge of language, whereas performance referred to actual language use under concrete conditions. The theoretical object was therefore not directly observable in itself; it had to be inferred from performance and other evidence. This distinction is important for educational evaluation because competencies, likewise, are not 'seen' directly but inferred from performances elicited by appropriately designed tasks.

Hymes extended this linguistic model by arguing that communicative competence depends not only on formal knowledge of language but also on the ability to use it appropriately in social contexts. This move brought contextualized use and observable performance into the conception of competence. For educational evaluation, the implication is important: claims about competence require evidence from performance in situations that make relevant resources and constraints visible, rather than from decontextualized knowledge alone.⁷

After this expansion in linguistics, the concept was developed further in professional and educational fields. Le Boterf defines competence in terms of acting appropriately by mobilizing relevant resources.⁸ Perrenoud similarly describes competence as the capacity to act effectively in a defined type of situation, drawing on knowledge without being reducible to it.⁹ This makes the relation between knowledge and competency explicit: knowledge is indispensable, but competence involves selecting,

⁵ Abdelrahman Toumi, *Competencies: A Systemic Approach*, Al-Hilal Publications, 5th ed., Morocco, 2006, p. 116.

⁶ Françoise Raynal and Alain Rieunier, *Pédagogie: Dictionnaire des concepts clés*, 7th ed., ESF, Paris, 2009, p. 401.

⁷ Noam Chomsky, *Aspects of the Theory of Syntax*, MIT Press, Cambridge, MA, 1965; Dell Hymes, "On Communicative Competence," in J. B. Pride and J. Holmes (eds.), *Sociolinguistics: Selected Readings*, Penguin, Harmondsworth, 1972, pp. 269-293.

⁸ Guy Le Boterf, *De la compétence à la navigation professionnelle*, Éditions d'Organisation, Paris, 1997, p. 42.

⁹ Philippe Perrenoud, *Construire des compétences dès l'école*, ESF, Paris, 1997.

coordinating, and mobilizing resources in response to a situation. Roegiers adds the idea of integrated mobilization across a family of situations.¹⁰ Accordingly, the object of evaluation becomes not the possession of isolated resources alone but the quality and appropriateness of their mobilization in context.

These developments show that the transformation affects the object of evaluation. Traditional measurement may validly describe discrete knowledge or skills, but a claim about competence is broader: it concerns whether the learner can integrate relevant resources and act effectively in a class of situations. Because competence is an inference from performance, the quality of that inference depends on the tasks used to elicit evidence and the criteria used to interpret it.

A further implication concerns transfer. Competency claims normally extend beyond the exact task on which performance is observed. When a learner succeeds in one integration situation, the evaluator is implicitly making a broader claim that the learner can mobilize relevant resources in other situations belonging to the same family. This is why a single successful performance should be interpreted cautiously. Competency is better supported by converging evidence across tasks that vary in surface features while preserving the essential demands of the construct. Such sampling reduces the risk of confusing familiarity with genuine mastery. It also highlights the importance of distinguishing resources from their mobilization: a learner may possess the required knowledge and procedures but fail to identify when or how to use them, while another learner may reach an acceptable solution through partial or inefficient mobilization. Evaluation must therefore attend not only to the final answer but, when appropriate, to strategy, justification, adaptation, and the quality of coordination among resources. This makes competency evaluation analytically richer than simple recall testing without making discrete knowledge assessment unnecessary.

This leads directly to the question of assessment tools.

From Test Questions to the Evaluation Situation:

As competency constructs became broader, isolated achievement-test items were no longer sufficient by themselves to support strong inferences about integrated performance. This does not mean that tests are obsolete: well-designed items can assess complex reasoning and provide useful measurement evidence. The limitation arises when a narrow item format is used as the sole basis for judging a construct that requires contextualized mobilization of multiple resources. Competency assessment therefore calls for a diversified evidence base, including tasks that approximate the situations in which the competency is expected to operate.

In the pedagogy of integration, Roegiers proposes the integration situation as one such tool. An integration situation is a complex task in which learners must reinvest and combine previously acquired resources in order to address a meaningful problem.¹¹ Its evaluative value depends not only on complexity but also on alignment with the targeted competency and on explicit criteria for judging performance. An evaluative integration situation should be:

- **Complex:** It requires the mobilization of several relevant resources (knowledge, skills, strategies, attitudes) rather than the reproduction of a single isolated element.
- **Meaningful:** It is situated in a context that is intelligible and relevant to the learner and gives the task a clear purpose.
- **New but related:** It is not a previously memorized task, yet it belongs to the same family of situations for which the competency has been developed. This makes it possible to judge transfer and adaptive mobilization rather than simple recall.

These conditions help make performance evidence more representative of the targeted competency. Yet the form of the task alone does not guarantee valid evaluation. The task must elicit the intended construct, scoring criteria must be sufficiently explicit, and judgments should be based on evidence that is adequate for the decision being made. Integration situations should therefore complement, rather than mechanically replace, other sources of evidence across the teaching-learning process.

The use of complex situations also raises the question of criteria. If competency performance is judged globally and without explicit descriptors, evaluation may become highly dependent on the individual evaluator. For this reason, complex tasks should be accompanied by criteria that specify the qualities expected in the performance, such as relevance of selected resources, accuracy, coherence, completeness, and appropriateness to the situation. Indicators can then make these criteria observable without fragmenting the competency into unrelated micro-skills. Analytic scoring may be useful when diagnostic feedback is required, whereas a more holistic judgment may be appropriate when the purpose is to determine the overall adequacy of an integrated performance. In both cases, the criteria should be communicated clearly and applied consistently. This strengthens transparency for learners, supports feedback, and makes professional judgment more defensible. It also

¹⁰ Xavier Roegiers, *La pédagogie de l'intégration: Des systèmes d'éducation et de formation au cœur de nos sociétés*, De Boeck Supérieur, Brussels, 2010, p. 66.

¹¹ Xavier Roegiers, *Des situations pour intégrer les acquis scolaires*, 2nd ed., De Boeck Supérieur, Louvain-la-Neuve, 2007, p. 38.

shows that moving from test questions to evaluation situations is not merely a change of format; it is a change in the architecture of evidence, interpretation, and decision-making.

From the foregoing, three major conceptual transformations can be identified:

- Transformation in the object: The focus expands from the possession or recall of discrete knowledge to the integrated mobilization of knowledge, skills, strategies, and attitudes in context.
- Transformation in the tool: Evaluation expands from reliance on isolated questions toward a diversified set of evidence sources, including well-designed tests, observations, projects, and integration situations. The tool should be selected according to the competency claim and the decision to be supported.
- Transformation in function: Evaluation shifts from a predominantly certification-oriented endpoint toward a process with diagnostic, formative, and summative functions across the learning sequence. The same evidence may serve different purposes depending on how it is interpreted and used.

A Contemporary Synthesis: Measurement, Evidence, and Judgment

Contemporary assessment theory reinforces an integrative rather than oppositional view. The National Research Council's 'assessment triangle' links three elements: a model of cognition, observations or tasks that elicit evidence, and interpretation procedures that connect that evidence to claims about learning.¹² Applied to competencies, this means that a complex task is useful only when it is aligned with a clear model of the competency and produces evidence that can be interpreted according to defensible criteria.

Classroom assessment research reaches a compatible conclusion. Black and Wiliam place assessment within pedagogy and emphasize the formative use of evidence to adapt teaching and learning.¹³ Bennett, however, cautions that formative purposes do not remove the need for sound measurement and domain-relevant evidence.¹⁴ Together, these perspectives show that quantitative measurement, qualitative interpretation, feedback, and pedagogical decision-making are complementary rather than mutually exclusive.

Recent systematic work also shows that competence-based education is not a single uniform model; its meanings and curriculum forms vary across systems and policy contexts.¹⁵ This makes construct clarity especially important. The key question is not whether a task is labelled 'competency-based,' but whether the targeted competency, the eliciting situation, the scoring criteria, and the interpretation of evidence are coherent with one another. Validity therefore becomes the bridge between the theoretical construct of competency and the practical act of evaluation.

The contemporary perspective also brings reliability, fairness, and assessment literacy into the discussion. Because competency judgments often rely on extended performances, projects, observations, or complex tasks, the evaluator must manage sources of variation that are less visible in short objective tests. Task difficulty, contextual familiarity, language demands, scoring criteria, and rater expectations can all influence the evidence produced. Fair evaluation therefore requires alignment between the intended competency and the demands that are actually necessary for demonstrating it. Irrelevant obstacles should not determine the judgment, and learners should have adequate opportunities to demonstrate the targeted competency. At the same time, teachers need the capacity to design tasks, formulate criteria, interpret evidence, provide feedback, and distinguish formative decisions from certification decisions. Assessment literacy is thus not a technical addition to competency-based education; it is one of the conditions for implementing it coherently. Without it, the vocabulary of competencies may change while classroom evaluation remains dominated by traditional scoring practices.

Conclusion

The analysis shows that the movement from measuring knowledge to evaluating competencies is best understood as an expansion and reorganization rather than a complete epistemological rupture. Measurement quantifies selected attributes; assessment elicits and interprets evidence; evaluation uses that evidence, together with criteria and judgment, to make pedagogical decisions. In competency-based education, knowledge remains indispensable, but it functions as one resource among others that must be mobilized in context. Consequently, the object, tools, and functions of evaluation broaden: from

¹² National Research Council, *Knowing What Students Know: The Science and Design of Educational Assessment*, National Academies Press, Washington, DC, 2001, doi:10.17226/10019.

¹³ Paul Black and Dylan Wiliam, "Classroom Assessment and Pedagogy," *Assessment in Education: Principles, Policy & Practice* 25, no. 6 (2018), pp. 551-575, doi:10.1080/0969594X.2018.1441807.

¹⁴ Randy Elliot Bennett, "Formative Assessment: A Critical Review," *Assessment in Education: Principles, Policy & Practice* 18, no. 1 (2011), pp. 5-25, doi:10.1080/0969594X.2010.513678.

¹⁵ Armend Tahirsylaj and Daniel Sundberg, "Five Visions of Competence-Based Education and Curricula as Travelling Policies: A Systematic Research Review 1997-2022," *Journal of Curriculum Studies* 58, no. 1 (2026), pp. 1-26, doi:10.1080/00220272.2025.2492605.

discrete recall to integrated performance; from reliance on a single item format to multiple sources of evidence; and from end-point certification to diagnostic, formative, and summative uses across the learning process.

The practical implication is that competency-based evaluation requires explicit criteria, task validity, transparent scoring, feedback, and teacher assessment literacy. The most defensible model is therefore neither "measurement versus evaluation" nor "knowledge versus competency," but a nested relationship in which measurement contributes evidence to broader assessment and evaluation processes. Future empirical research should examine how this conceptual model is implemented in classroom practices and how reliably integration situations support judgments about learners' competencies.

References

1. Abdelaziz Quraish. "Educational Evaluation: The Question of Nature and Function." *Education and Training Notebooks*, no. 4, February 2011, Morocco.
2. Abdelrahman Toumi. *Competencies: A Systemic Approach*. 5th ed. Al-Hilal Publications, Morocco, 2006.
3. Bennett, R. E. "Formative Assessment: A Critical Review." *Assessment in Education: Principles, Policy & Practice* 18, no. 1 (2011): 5-25. doi:10.1080/0969594X.2010.513678.
4. Black, P., and D. Wiliam. "Classroom Assessment and Pedagogy." *Assessment in Education: Principles, Policy & Practice* 25, no. 6 (2018): 551-575. doi:10.1080/0969594X.2018.1441807.
5. Chomsky, N. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press, 1965.
6. de Landsheere, G. *Dictionnaire de l'évaluation et de la recherche en éducation*. 2nd rev. and expanded ed. Paris: PUF, 1992.
7. de Landsheere, G. *Évaluation continue et examens: Précis de docimologie*. 5th rev. ed. Brussels: Labor, 1980.
8. Hymes, D. H. "On Communicative Competence." In *Sociolinguistics: Selected Readings*, edited by J. B. Pride and J. Holmes, 269-293. Harmondsworth: Penguin, 1972.
9. Le Boterf, G. *De la compétence à la navigation professionnelle*. Paris: Éditions d'Organisation, 1997.
10. National Research Council. *Knowing What Students Know: The Science and Design of Educational Assessment*. Washington, DC: National Academies Press, 2001. doi:10.17226/10019.
11. Perrenoud, P. *Construire des compétences dès l'école*. Paris: ESF, 1997.
12. Raynal, F., and A. Rieunier. *Pédagogie: Dictionnaire des concepts clés*. 7th ed. Paris: ESF, 2009.
13. Roegiers, X. *Des situations pour intégrer les acquis scolaires*. 2nd ed. Louvain-la-Neuve: De Boeck Supérieur, 2007. doi:10.3917/dbu.roegi.2007.02.
14. Roegiers, X. *La pédagogie de l'intégration: Des systèmes d'éducation et de formation au cœur de nos sociétés*. Brussels: De Boeck Supérieur, 2010.
15. Stevens, S. S. "On the Theory of Scales of Measurement." *Science* 103, no. 2684 (1946): 677-680. doi:10.1126/science.103.2684.677.
16. Tahirsylaj, A., and D. Sundberg. "Five Visions of Competence-Based Education and Curricula as Travelling Policies: A Systematic Research Review 1997-2022." *Journal of Curriculum Studies* 58, no. 1 (2026): 1-26. doi:10.1080/00220272.2025.2492605.