CRITERIA FOR ASSESSING WRITTEN COMPETENCE IN PRE-SERVICE ENGLISH TEACHERS: A RUBRIC-BASED FORMATIVE AND SUMMATIVE ASSESSMENT FRAMEWORK ALIGNED WITH THE CEFR

Guzal Turaeva

Namangan State University, Doctoral student
E-mail:guzalturaeva025@gmail.com

ARTICLE INFORMATION	ABSTRACT:
ARTICLE HISTORY: *Received:29.01.2026* *Revised: 30.01.2026* *Accepted:31.01.2026* **KEYWORDS:** *written competence; writing assessment; rubric-based assessment; formative assessment; summative assessment; educational evaluation*	*This paper presents an empirical synthesis of peer-reviewed research examining criteria for assessing students' written competence within rubric-based formative and summative assessment frameworks. Drawing on systematic reviews, meta-analyses, and empirical studies in educational assessment and writing research, the paper identifies core, evidence-based criteria commonly used to operationalize written competence, including task fulfillment, organization and coherence, content development, language use, vocabulary, and mechanics. The synthesis further examines how these criteria are embedded in analytic rubrics and how their use differs in formative versus summative assessment contexts. Findings indicate that rubrics are most effective when criteria are clearly defined, aligned with instructional goals, and integrated into formative feedback cycles that support revision and self-regulation. In summative contexts, rubric effectiveness is strongly associated with rater training, calibration procedures, and alignment with instructional practices, which collectively enhance reliability and validity. The paper argues for a*

==

hybrid assessment approach in which the same rubric-based criteria support both learning-oriented formative assessment and accountability-driven summative assessment. By aligning assessment criteria with instructional practices, educators can improve the fairness, transparency, and educational impact of writing assessment. The study contributes to assessment theory by clarifying how written competence can be systematically and meaningfully evaluated in educational settings and offers practical implications for educators, researchers, and assessment designers

Introduction. Writing is a central mode of communication and a primary means through which students demonstrate knowledge, reasoning, and disciplinary understanding across educational contexts. As such, the assessment of students' written competence plays a crucial role in shaping instructional practices, learning outcomes, and academic progression. Despite its importance, writing assessment remains one of the most contested and methodologically complex areas in educational evaluation due to the multifaceted nature of writing and the interpretive judgment involved in its assessment.

Written competence encompasses more than grammatical accuracy or surface correctness. It involves the ability to generate ideas, structure arguments, adapt language to audience and purpose, and communicate meaning effectively within specific genres and contexts. Consequently, assessing writing requires evaluative frameworks capable of capturing this complexity without reducing writing to isolated linguistic features. Traditional approaches to writing assessment have often been criticized for their subjectivity, lack of transparency, and limited instructional value. In response, rubric-based assessment has gained prominence as a tool designed to clarify expectations, standardize evaluation, and support feedback. At the same time, assessment theory has increasingly emphasized the distinction and interaction between formative assessment, which aims to support learning during instruction, and summative assessment, which aims to evaluate learning outcomes at specific points in time (Black & Wiliam, 1998).

In response to these challenges, rubric-based assessment has become a widely adopted approach in both school and higher education contexts. Rubrics aim to articulate explicit criteria for quality writing, enhance consistency in scoring, and support feedback processes. Simultaneously, contemporary assessment theory emphasizes the integration of formative assessment-assessment for learning, and summative assessment-assessment of learning-as complementary rather than opposing functions (Black & Wiliam, 1998).

This chapter examines criteria for assessing students' written competence through the lens of rubric-based assessment within formative and summative frameworks. Drawing on empirical studies, systematic reviews, and meta-analyses, the chapter synthesizes evidence on effective rubric design, the validity and reliability of rubric-based assessment, and the pedagogical implications of integrating rubrics into writing instruction.

The chapter pursues three main objectives:

1. To conceptualize written competence and identify empirically supported assessment criteria;

2. To examine the role of rubrics in formative and summative assessment of writing;

3. To provide evidence-based implications for writing assessment practices in educational contexts.

Literature Review

Writing assessment research. Research on writing assessment has evolved alongside shifts in writing pedagogy. Early product-oriented approaches focused primarily on correctness and surface features, while process-oriented perspectives emphasized drafting, revision, and cognitive strategies (Graham & Perin, 2007). More recent sociocognitive and genre-based perspectives highlight the role of audience, purpose, and disciplinary conventions in shaping writing performance. Empirical studies consistently demonstrate that narrow assessment criteria focused solely on mechanics fail to capture higher-order writing skills and may distort instructional priorities. Conversely, multidimensional assessment frameworks that integrate content, organization, language use, and rhetorical effectiveness provide more valid representations of written competence (Weigle, 2002).

Rubrics in educational assessment. Rubrics have been widely studied as tools for both instruction and evaluation. Reddy and Andrade (2010), in a comprehensive review, found that rubrics can improve transparency, support self-regulated learning, and enhance scoring consistency when well designed. However, they caution that rubrics are not inherently effective; their impact depends on clarity of criteria, alignment with instruction, and the

==

manner in which they are used. Research also highlights potential drawbacks, including over-standardization and mechanical application, which can constrain creativity or reduce writing to checklist compliance if rubrics are poorly implemented.

Conceptual Framework of Written Competence

Defining written competence. Written competence is conceptualized in this chapter as a multidimensional construct integrating cognitive, linguistic, and rhetorical components. This framework draws on models of writing that emphasize interaction between planning, text production, and revision processes, mediated by language knowledge and contextual awareness (Graham & Perin, 2007).

At the cognitive level, written competence involves idea generation, organization, and critical thinking. At the linguistic level, it encompasses grammatical accuracy, lexical choice, and syntactic complexity. At the rhetorical level, it includes audience awareness, genre conventions, and communicative purpose. Effective assessment must therefore address these dimensions through clearly articulated criteria that reflect instructional goals and disciplinary expectations..From an assessment perspective, written competence must be inferred from observable textual features. These features function as indicators of underlying abilities but cannot capture the full complexity of the writing process. As a result, assessment frameworks typically operationalize written competence through multiple criteria, each representing a distinct dimension of writing quality (Weigle, 2002). Rubrics serve as mediating tools within this framework, translating abstract constructs of writing quality into observable and assessable descriptors.

Writing as a construct in assessment. Writing differs from many other academic skills because it requires the integration of higher-order thinking skills and lower-level linguistic skills. Empirical research demonstrates that focusing exclusively on surface features such as grammar and spelling leads to narrow interpretations of competence and may disadvantage students who demonstrate strong ideas but weaker linguistic control (Graham & Perin, 2007).

Therefore, contemporary assessment models emphasize a balance between content-related criteria (e.g., argument development, organization) and language-related criteria (e.g., grammatical accuracy, vocabulary use). This balanced view underpins rubric-based assessment systems used in both classroom and large-scale contexts.

==

==

Theoretical Perspectives on Assessment

Formative assessment. Formative assessment is defined as a set of practices that provide information used by teachers and learners to adjust instruction and learning strategies. Black and Wiliam (1998) demonstrated through a large-scale review that formative assessment has a substantial positive impact on student achievement across subjects, including writing. A central feature of formative assessment is feedback. Sadler (1989) emphasizes that formative assessment is effective only when learners understand quality criteria, can evaluate their own performance, and are supported in closing performance gaps. In writing instruction, rubrics can facilitate these conditions by clarifying expectations and guiding revision.

Summative assessment. Summative assessment serves the purpose of certifying learning outcomes, often for grading or accountability. In writing assessment, summative evaluation raises concerns regarding fairness, reliability, and transparency, particularly because scoring depends on human judgment. Empirical research emphasizes that summative writing assessment must rely on clearly articulated criteria, standardized procedures, and rater training to minimize bias and inconsistency (Jonsson & Svingby, 2007). While summative assessment is often perceived as incompatible with formative goals, recent scholarship suggests that alignment between formative and summative practices enhances validity and instructional coherence.

Rubrics as Instruments for Writing Assessment

Definition and purpose of rubrics. A rubric is a scoring guide that articulates assessment criteria and describes levels of performance for each criterion. Rubrics aim to make assessment expectations explicit, support consistent scoring, and provide meaningful feedback to learners. In writing assessment, rubrics function as both evaluative and instructional tools. When used effectively, they can support learning by clarifying what constitutes quality writing and guiding revision processes (Reddy & Andrade, 2010).

Types of rubrics. Rubrics are commonly categorized as holistic or analytic. Holistic rubrics assign a single overall score based on an overall impression of writing quality, whereas analytic rubrics separate writing into multiple criteria, each scored independently. Empirical evidence suggests that analytic rubrics are more suitable for instructional contexts because they provide detailed information about strengths and weaknesses across dimensions of writing (Jonsson & Svingby, 2007). Analytic rubrics also facilitate targeted feedback and self-assessment, making them particularly effective for formative assessment.

==

==

By contrast, holistic rubrics may be appropriate in large-scale or time-constrained contexts where efficiency is prioritized.

Reliability and Validity of Rubric-Based Writing Assessment

Reliability in writing assessment. Reliability refers to the consistency of assessment results across raters or occasions. Writing assessment is particularly vulnerable to reliability issues due to subjective judgment. Jonsson and Svingby (2007) found that rubrics can improve interrater reliability when criteria are explicit and raters receive training and calibration opportunities. However, reliability is not inherent in the rubric itself; it depends on how the rubric is used. Studies show that poorly defined criteria or inconsistent interpretation of descriptors can undermine reliability even in rubric-based systems.

Validity considerations. Validity concerns whether an assessment measures the intended construct. In writing assessment, validity requires that criteria accurately represent the dimensions of written competence valued in instruction and curriculum. Overemphasis on easily measurable features, such as mechanics, can threaten construct validity by marginalizing higher-order skills. Empirical research highlights the importance of aligning rubric criteria with instructional goals and writing tasks to ensure meaningful interpretation of scores (Weigle, 2002).

Empirically Supported Criteria for Assessing Written Competence

Task fulfillment and communicative purpose. Task fulfillment refers to the extent to which a text addresses the prompt and achieves its intended purpose. Empirical studies consistently identify task fulfillment as a core criterion because writing competence is context-dependent and goal-oriented. Effective rubrics articulate task fulfillment in terms of audience awareness, genre conventions, and communicative intent, rather than mere topic relevance.

Organization and coherence. Organization and coherence are central to reader comprehension. Research on writing instruction indicates that explicit teaching and assessment of organizational strategies significantly improve students' writing quality (Graham & Perin, 2007). Rubric descriptors typically address logical sequencing, paragraph structure, and the use of cohesive devices. Empirical evidence supports treating organization as a distinct criterion rather than subsuming it under content or language use.

Content development and argumentation. Content development refers to the depth, relevance, and support of ideas. In analytical and argumentative writing, this includes the use of evidence, reasoning, and critical engagement with the topic. Empirical studies show

==

that rubrics emphasizing content development encourage students to move beyond surface-level responses and engage in higher-order thinking.

Language use and grammatical control. Language use encompasses sentence structure, grammatical accuracy, and clarity of expression. Research indicates that grammar-focused feedback is most effective when integrated with meaning-focused instruction rather than treated as an isolated skill (Graham & Perin, 2007). Consequently, effective rubrics evaluate grammatical accuracy in relation to communicative effectiveness rather than error frequency alone.

Vocabulary and stylistic appropriateness. Vocabulary use contributes to precision and tone. Empirical research suggests that assessing vocabulary in terms of appropriateness and effectiveness is more valid than focusing on lexical complexity alone.

Mechanics and presentation. Mechanics, including spelling, punctuation, and formatting, influence readability and adherence to academic conventions. Empirical literature supports assigning lower weighting to mechanics to avoid overshadowing higher-order writing skills.

Rubrics in Formative Assessment of Writing

Rubrics as learning tools. When used formatively, rubrics function as tools for learning rather than merely instruments of judgment. Reddy and Andrade (2010) report that students who use rubrics for self- and peer assessment demonstrate improved understanding of quality writing and greater self-regulation.

Feedback and revision cycles. Empirical evidence underscores the importance of integrating rubrics into feedback cycles. Effective formative use involves introducing rubrics before writing tasks, encouraging self-assessment, and providing criterion-referenced feedback during drafting and revision. However, research cautions against overloading rubrics with excessive criteria, as this may reduce their instructional effectiveness.

Rubrics in Summative Assessment of Writing

Scoring consistency and fairness. In summative contexts, rubrics primarily function to support consistent and fair scoring. Empirical studies highlight the role of rater training, calibration sessions, and moderation procedures in enhancing reliability (Jonsson & Svingby, 2007).

Alignment with instruction. Summative rubrics are most valid when aligned with instructional practices and formative assessment criteria. Misalignment can undermine students' perceptions of fairness and reduce the educational value of assessment outcomes.

Methodological Orientation: Empirical Synthesis. This chapter adopts an empirical synthesis approach, integrating findings from peer-reviewed studies, systematic reviews, and meta-analyses. This approach allows for the identification of recurring patterns and convergent evidence related to writing assessment criteria and rubric use. Rather than reporting original experimental data, the synthesis draws conclusions based on the collective weight of empirical research, which is appropriate for thesis chapters focused on theoretical and methodological foundations.

Discussion. The purpose of this chapter is to examine criteria for assessing students' written competence and to explore how rubric-based assessment functions within formative and summative assessment frameworks. Drawing on empirical research and established assessment theory, the discussion interprets the synthesized findings in relation to broader debates in writing assessment and educational evaluation. This section revisits the central themes of the chapter, highlights theoretical and practical tensions, and clarifies the contribution of rubric-based assessment to understanding and evaluating written competence.

Written competence as a multidimensional construct. One of the central insights emerging from this synthesis is the necessity of conceptualizing written competence as a multidimensional construct rather than a unitary skill. Across the reviewed literature, writing competence is consistently operationalized through multiple criteria, including task fulfillment, organization, content development, language use, vocabulary, and mechanics (Weigle, 2002; Graham & Perin, 2007). This finding reinforces arguments that reductive approaches to writing assessment—particularly those privileging grammatical accuracy alone—fail to represent the full scope of writing ability.

The discussion aligns with sociocognitive and genre-based perspectives on writing, which emphasize that writing quality is context-sensitive and shaped by audience, purpose, and disciplinary norms (Hyland, 2016). Rubric-based assessment, when grounded in clearly articulated criteria, provides a mechanism for translating these complex theoretical constructs into assessable dimensions without oversimplifying the writing task.

==

Rubrics as mediating tools between theory and practice. A key contribution of this chapter lies in framing rubrics not merely as scoring instruments, but as mediating tools that connect assessment theory with instructional practice. From a formative assessment perspective, rubrics operationalize standards of quality and make evaluative criteria visible to learners, thereby supporting self-assessment and revision (Sadler, 1989; Panadero & Jonsson, 2013). The synthesis indicates that rubrics are most effective when they function dialogically—supporting feedback exchanges between teachers and students—rather than as static checklists. This interpretation aligns with assessment-for-learning frameworks, which emphasize learner engagement with criteria as a prerequisite for meaningful feedback (Black & Wiliam, 1998). Thus, the educational value of rubrics depends less on their formal structure and more on how they are embedded in pedagogical practices.

Tensions between formative and summative purposes. The findings also illuminate enduring tensions between formative and summative assessment purposes in writing evaluation. While formative assessment prioritizes learning, feedback, and development, summative assessment emphasizes reliability, comparability, and accountability. The literature reviewed in this chapter demonstrates that rubrics are often expected to serve both purposes simultaneously, which can lead to conflicting demands (Jonsson & Svingby, 2007).

This discussion supports integrative assessment models that view formative and summative assessment as complementary rather than dichotomous. When shared criteria guide instruction, formative feedback, and summative judgment, assessment practices are more coherent and valid. However, the synthesis also cautions that such integration requires careful design, particularly in high-stakes contexts where summative pressures may dominate and undermine formative intentions.

Reliability, validity, and the limits of standardization. Another important theme concerns the relationship between reliability and validity in rubric-based writing assessment. Empirical studies consistently show that rubrics can enhance interrater reliability when criteria are explicit and raters are trained (Jonsson & Svingby, 2007). Nevertheless, increased reliability does not automatically guarantee validity. Overly rigid or narrowly defined criteria may improve scoring consistency at the expense of construct representation. This discussion reinforces theoretical critiques of excessive standardization in writing assessment, which warn against reducing writing competence to easily measurable features (Weigle, 2002). The findings suggest that valid assessment requires a

==

balance between standardization and professional judgment, with rubrics serving as guides rather than substitutes for informed evaluation.

Student engagement and assessment literacy. The synthesis highlights the role of rubrics in developing students' assessment literacy—their ability to understand, interpret, and apply assessment criteria. Research reviewed in this chapter indicates that when students actively engage with rubrics through self- and peer assessment, they develop greater metacognitive awareness and control over their writing processes (Panadero & Jonsson, 2013; Reddy & Andrade, 2010). From a theoretical standpoint, this finding supports models of self-regulated learning, which position assessment as a driver of learner autonomy. The discussion suggests that rubric-based assessment contributes not only to more transparent evaluation, but also to long-term writing development when learners internalize criteria for quality writing.

Implications for conceptualizing writing assessment. Taken together, the findings discussed in this chapter challenge narrow, product-oriented conceptions of writing assessment. Instead, they support a view of assessment as an ongoing, interpretive, and pedagogically embedded practice. Rubrics, when thoughtfully designed and used, provide a flexible framework capable of accommodating the complexity of written competence while addressing institutional demands for fairness and consistency. Importantly, this discussion underscores that rubrics are not neutral artifacts; they embody particular values about what counts as quality writing. Making these values explicit through transparent criteria and reflective use is essential for equitable and meaningful assessment practices, particularly in diverse educational contexts.

Implications for Educational Practice

Based on empirical evidence, educators are encouraged to:

- Design analytic rubrics aligned with learning objectives;
- Use rubrics formatively to support revision and self-assessment;
- Provide rater training for summative assessment;
- Limit the number of criteria to essential dimensions of writing competence;
- Engage students actively in understanding assessment standards.

Limitations and Directions for Future Research. Despite providing a comprehensive empirical synthesis of research on rubric-based assessment and criteria for assessing students' written competence, this chapter is subject to several limitations that should be

===

acknowledged. Recognizing these limitations is essential for interpreting the findings appropriately and for guiding future research in writing assessment.

Methodological limitations. The primary limitation of this chapter lies in its reliance on an empirical synthesis methodology rather than the collection of original primary data. While empirical synthesis allows for the identification of convergent evidence across multiple studies, it is inherently constrained by the scope, quality, and methodological diversity of the existing literature (Cooper, 2017). Consequently, the conclusions drawn in this chapter reflect prevailing research trends rather than causal claims derived from experimental investigation. Moreover, the reviewed studies employ varied research designs, including quasi-experimental studies, case studies, and systematic reviews. This heterogeneity strengthens the breadth of the synthesis but limits the ability to generalize findings across all educational contexts or learner populations (Gough et al., 2012).

Contextual and disciplinary limitations. A further limitation concerns the contextual specificity of writing assessment research. Much of the empirical literature on rubric-based writing assessment has been conducted in Western educational systems, particularly in English-dominant contexts (Weigle, 2002; Jonsson & Svingby, 2007). As writing is a socially and culturally situated practice, assessment criteria and interpretations of writing quality may vary significantly across cultural, linguistic, and institutional settings. In addition, writing conventions differ across academic disciplines. Criteria commonly used to assess argumentative writing in education or the humanities may not fully capture writing competence in scientific, technical, or professional disciplines (Lea & Street, 1998). This disciplinary variation presents challenges for the development of universally applicable rubrics and underscores the need for context-sensitive assessment frameworks.

Directions for future research. Future research should address these limitations through several complementary avenues.

- First, context-specific empirical studies are needed to examine how rubric-based assessment functions across diverse educational, linguistic, and cultural contexts. Such research would enhance understanding of how written competence is locally constructed and evaluated (Hyland, 2016).

- Second, there is a clear need for longitudinal studies investigating the sustained impact of rubric-based formative assessment on students' writing development. While existing research provides evidence of short-term learning gains, less is known about how prolonged

===

engagement with rubrics influences writing proficiency, metacognitive awareness, and self-regulated learning over time (Panadero & Jonsson, 2013).

- Third, future studies should explore discipline-specific writing assessment, focusing on how rubrics can be adapted to reflect the epistemological and rhetorical norms of different academic fields. Such research would strengthen the construct validity of writing assessment in higher education (Lea & Street, 1998).

Finally, emerging digital and automated writing assessment technologies present new opportunities and challenges. Future research should investigate how digital rubrics, automated feedback systems, and artificial intelligence tools interact with human judgment and pedagogical practices, particularly in relation to reliability, validity, and equity (Dikli, 2006).

Conclusion. The assessment of students' written competence remains one of the most complex and consequential challenges in educational assessment. Writing is a multidimensional construct that integrates cognitive processes, linguistic resources, and rhetorical awareness, making it resistant to simplistic or purely mechanistic evaluation (Weigle, 2002). This chapter has argued that effective writing assessment requires a principled balance between technical measurement standards and pedagogical responsiveness. Drawing on empirical research and assessment theory, this chapter has demonstrated that rubric-based assessment provides a robust framework for operationalizing written competence when criteria are clearly defined, aligned with instructional goals, and grounded in authentic writing practices (Jonsson & Svingby, 2007; Reddy & Andrade, 2010). Analytic rubrics, in particular, allow assessors to capture multiple dimensions of writing quality while supporting transparency, consistency, and meaningful feedback.

References

1. Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. Phi Delta Kappan, 80(2), 139–148.

2. Cooper, H. (2017). Research synthesis and meta-analysis: A step-by-step approach (5th ed.). SAGE.

3. Dikli, S. (2006). An overview of automated scoring of essays. Journal of Technology, Learning, and Assessment, 5(1), 1–35.

==

4. Graham, S., & Perin, D. (2007). A meta-analysis of writing instruction for adolescent students. Journal of Educational Psychology, 99(3), 445–476. https://doi.org/10.1037/0022-0663.99.3.445

5. Gough, D., Oliver, S., & Thomas, J. (2012). An introduction to systematic reviews. SAGE.

6. Hyland, K. (2016). Teaching and researching writing (3rd ed.). Routledge.

7. Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. Educational Research Review, 2(2), 130–144. https://doi.org/10.1016/j.edurev.2007.05.002

8. Lea, M. R., & Street, B. V. (1998). Student writing in higher education: An academic literacies approach. Studies in Higher Education, 23(2), 157–172.

9. Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited. Educational Research Review, 9, 129–144.

10. Reddy, Y. M., & Andrade, H. (2010). A review of rubric use in higher education. Assessment & Evaluation in Higher Education, 35(4), 435–448. https://doi.org/10.1080/02602930902862859

11. Sadler, D. R. (1989). Formative assessment and the design of instructional systems. Instructional Science, 18(2), 119–144. https://doi.org/10.1007/BF00117714

12. Weigle, S. C. (2002). Assessing writing. Cambridge University Press.

==