Contents lists available at Science-Gate

International Journal of Advanced and Applied Sciences

Journal homepage: https://www.science-gate.com/IJAAS.html

Proposing writing competency assessment standards for upper secondary students in Vietnam

Ha Do Thu [1], Hao Thi Nguyen [2, *], Nga Thi Thanh Nguyen [3], Lan Thi Huong Nguyen [4], Van Hong Thi Nguyen [5], Diep Thi Be [6]

[1]*Center for Education Policy Research, The Vietnam National Institute of Education Sciences, Hanoi, Vietnam*
[2]*Center for Human Resource Development Research, The Vietnam National Institute of Educational Sciences, Hanoi, Vietnam*
[3]*Center for Curriculum Development and Educational Assessment Research, The Vietnam National Institute of Educational Sciences, Hanoi, Vietnam*
[4]*Office of Science Management and Training, The Vietnam National Institute of Educational Sciences, Hanoi, Vietnam*
[5]*The Vietnam National Institute of Educational Sciences, Hanoi, Vietnam*
[6]*National Center for Educational Testing and Quality Assessment, Quality Management Department, Ministry of Education and Training, Hanoi, Vietnam*

ARTICLE INFO

Article history:
Received 22 August 2025
Received in revised form
10 July 2026
Accepted 29 September 2026

Keywords:
Competency-based assessment
Argumentative writing
Vietnamese literature
Secondary education
Writing skills

ABSTRACT

This study develops and validates a competency-based assessment standard for argumentative writing in Vietnamese literature at the upper secondary level, aligned with the 2018 General Education Curriculum of Vietnam. The study outlines a clear procedure for constructing the standard, including the identification of required skills, the development of assessment criteria and performance levels, and pilot testing in both rural and urban high schools. The proposed standard consists of four main components and twelve sub-criteria and was applied to a sample of 164 Grade 10 students. The results confirm high reliability and strong structural validity, as indicated by Cronbach's alpha (α = 0.92), exploratory factor analysis, and Rasch modeling. In addition, Guttman scale analysis shows that the standard effectively distinguishes students across different levels of proficiency. The findings indicate that most students achieved or exceeded the basic requirements for core writing skills, while creativity was identified as the most challenging aspect. Overall, the results support the feasibility, reliability, and educational value of the proposed assessment framework, providing a useful tool for teachers to assess and support students' writing development. Implications for wider application and future research on writing and related language skills are also discussed.

© 2026 The Authors. Published by IASE. This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Education in the 21st century has transitioned from content-based approaches to competency-based learning, a shift aligned with Delors's (1996) four pillars of education: learning to know, learning to do, learning to live together, and learning to be. Reflecting this global trend, the Ministry of Education and Training of Vietnam (MOET) introduced the General Education Curriculum in 2018, which includes subject-specific curricula and educational activities designed to cultivate both general and subject-specific competencies. Despite the inclusion of detailed learning outcomes, the curriculum lacks formal, operationalized competency assessment standards—particularly in writing—which hinders its effective implementation, especially under the "one curriculum, many textbooks" policy framework.

Writing is recognized as a core competency assessed in national graduation exams and represents a crucial benchmark for evaluating student learning outcomes. Therefore, establishing formal, measurable standards for writing competency is essential for guiding both teaching and assessment practices in Vietnamese upper secondary schools. Competency-based assessment standards specify what students should know and be able to do, using observable indicators and performance levels evaluated through authentic tasks (Phuong et al., 2023). Recent research has further developed the understanding and

* Corresponding Author.
Email Address: haont@vnies.edu.vn (H. T. Nguyen)
https://doi.org/10.21833/ijaas.2026.09.017
Corresponding author's ORCID profile:
https://orcid.org/0009-0000-3774-0561
2313-626X/© 2026 The Authors. Published by IASE.
This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0/)

implementation of competency-based education and assessment. In the Vietnamese context, Yen and Thao (2024) examined competency-based assessment practices and highlighted the importance of aligning assessment methods with clearly defined learning outcomes and students' demonstrated competencies. Similarly, Ponomariovienė et al. (2025) emphasized the importance of measurable performance criteria, continuous progress monitoring, and the assessment of students' ability to apply knowledge and skills in practical situations. Research on educational assessment has also demonstrated the value of developing measurable indicators of student performance, particularly in writing proficiency, analytical reasoning, and the application of evidence. Collectively, these studies reinforce the importance of establishing explicit competency criteria, observable performance indicators, and systematic assessment instruments.

Recent research has continued to examine the theoretical and practical dimensions of literacy education and competency-based assessment. In Vietnam, Ngo et al. (2024) reviewed developments in literacy education and educational reform, highlighting the competency-oriented direction of the Vietnamese curriculum and the importance of developing reading, writing, and analytical abilities. More recent studies have also emphasized the importance of clearly defined assessment criteria for evaluating argumentative writing, including the organization of ideas, logical reasoning, supporting evidence, and the formulation of counterarguments (Tang et al., 2025). These developments reinforce the need for systematic assessment frameworks that translate the objectives of Vietnamese Literature education into explicit performance indicators and measurable standards.

Internationally, curriculum frameworks and writing assessment systems provide established foundations for defining and evaluating writing competency. The Common Core State Standards in the United States (learning.ccsso.org), the Australian Curriculum Version 9.0 (australiancurriculum.edu.au), and South Korea's 2022 Revised National Curriculum (moe.go.kr) emphasize the development of coherent texts, effective organization, appropriate language use, and adaptation of writing to different purposes and audiences. Recent research further demonstrates the importance of multidimensional writing assessment frameworks incorporating content development, organization, coherence, vocabulary, language accuracy, and genre-specific conventions (Imsa-ard, 2025; Mojumdar, 2026). These international frameworks and research findings provide a valuable foundation for developing contextually appropriate writing competency standards in Vietnamese upper secondary education.

Taken together, this body of literature offers a foundational base for developing a structured, context-appropriate writing competency assessment framework for Vietnamese upper secondary students. This study consists of three main components: (1) the theoretical rationale, (2) the development process, and (3) a proposed set of writing competency assessment standards for the subject of Vietnamese Literature.

Assessing student competency is a developmental process that monitors and supports the progressive acquisition of knowledge and skills. Three complementary theoretical perspectives provide a foundation for this process. First, Vygotsky's (1978) Zone of Proximal Development (ZPD) emphasizes the importance of evaluating not only students' independent performance but also their potential for further development through guidance, interaction, and appropriate instructional support. Recent applications of dynamic assessment demonstrate how this perspective can be used to identify emerging competencies and guide targeted instruction (Poehner, 2025). Second, Rasch measurement, a specific model within the broader Item Response Theory (IRT) family, establishes a probabilistic relationship between student ability and item difficulty, allowing performance to be measured and compared along a common scale (Hope et al., 2025). Third, learning progression frameworks conceptualize competency development as a sequence of increasingly sophisticated levels of knowledge and performance. These frameworks support the definition of observable performance indicators and the systematic monitoring of students' progress toward established competency standards.

By integrating these frameworks, assessment standards can reflect the developmental trajectory of learners, align learning outcomes with observable behaviors, and support differentiated instruction based on students' evolving competency levels.

The General Education Curriculum for Vietnamese Literature serves as the official national framework guiding textbook development, instructional planning, and assessment. The curriculum emphasizes not only the development of general qualities and competencies but also two subject-specific competencies: language competence and literary competence. These are operationalized across four skill domains: reading, writing, speaking, and listening.

Writing competency assessment standards must be developed in alignment with the curriculum's specified learning outcomes. They should translate these objectives into measurable indicators and define progressive levels of achievement that reflect students' mastery of writing skills.

Effective standards for assessing writing competency must demonstrate both scientific validity and pedagogical relevance. Validity is ensured when standards align with curriculum outcomes and are expressed through clear, measurable indicators. Pedagogical relevance requires that standards be realistic, accessible, and reflective of actual student performance at the upper secondary level. Moreover, the standards must differentiate between varying levels of proficiency and be usable by both educators and learners for

instructional and self-assessment purposes. Recent reforms in Vietnamese secondary education have promoted formative and summative assessment practices, including the use of rubrics, checklists, and observation tools. While these reforms show promise, challenges persist. Many teachers struggle with competency-based assessment principles, and assessments often fail to align with curriculum-based outcomes.

Additionally, current descriptions of writing competencies are too abstract, limiting their classroom applicability. Given the significance of writing in high-stakes examinations, a standardized framework is necessary to ensure objectivity, reliability, and instructional relevance in writing assessments.

Educational assessment standards are developed to evaluate the extent to which teaching and learning meet expected outcomes. They facilitate objective evaluation, foster student-centered learning environments, and support systemic educational improvement. In this study, the process of developing writing competency standards draws upon theoretical insights from Vygotsky (1978) and consists of five steps:

- Identify writing competency: Define the concept and outline its components within the context of upper secondary education.
- Determine sub-competencies: Analyze the structural elements of writing relevant to students at this educational level.
- Define behavioral indicators: Develop observable, grade-specific indicators that describe expected learning outcomes.
- Develop quality criteria: Articulate progressive levels of proficiency for each indicator, forming a descriptive scale from basic to advanced levels.
- Pilot the standards: Apply the draft standards to a representative student sample, collect feedback, and revise accordingly to ensure applicability in real-world settings.

We define writing competency assessment standards as clear, measurable descriptions of students' performance in writing across progressively advanced levels. These standards include four main components. Based on the General Education Curriculum for Vietnamese Literature, core writing competency components include:

- Content: Accuracy, relevance, and depth of ideas presented.
- Organization: Coherence, logical structure, and progression of arguments.
- Expression and Argumentation: Clarity of voice, persuasiveness, and rhetorical strategies.
- Language Use: Grammar, vocabulary, sentence structure, and stylistic features.
- Text Format: Adherence to genre conventions and formatting rules.

This framework is supported by evaluation rubrics from the National Upper Secondary Graduation Examination (starting in 2025), which reinforces the importance of structure, clarity, genre, and creativity. By synthesizing curriculum guidelines, textbook frameworks, and assessment rubrics, we propose a set of writing competency standards with corresponding behavioral indicators. These are designed to guide teaching, learning, and evaluation in Vietnamese Literature at the upper secondary level.

The four parts and twelve behavioral indicators in the suggested writing skills structure can be expanded into performance levels (quality criteria) to create assessment standards for each type of writing in the Vietnamese Literature curriculum. The assessment standards are divided into three levels, called M3, M2, and M1. The assessment standards are categorized into three levels, labeled as M3, M2, and M1. Among these, M1 represents the baseline standard (basic proficiency), M2 reflects a proficient level, and M3 indicates advanced performance. In the following section, we present an illustrative proposal for writing competency assessment standards for upper secondary students, with a focus on argumentative writing—a core writing type in the Vietnamese Literature curriculum at the upper secondary level in Vietnam. The proposed writing competency assessment standards, including the four main criteria, their corresponding sub-criteria, and the three performance levels (M1, M2, and M3), are presented in Table 1.

Table 1: Writing competency standards for argumentative essays for upper secondary students

Criterion	Sub-criterion	M1 (basic)	M2 (proficient)	M3 (advanced)
Content	Task identification	Identifies issue partially; focus inconsistent	Identifies issue; mostly maintains focus	Clearly identifies issue; fully consistent focus
	Knowledge accuracy	Basic accuracy; limited depth	Generally accurate and relevant	Accurate, relevant, and in-depth
	Viewpoint	Basic, generally appropriate	Clear and appropriate	Clear, insightful, and well-developed
	Thesis and structure	Some arguments; weak organization	Key arguments; mostly logical	Well-structured, coherent arguments
Argument development	Evidence use	Limited examples; weak support	Relevant evidence; adequate analysis	Strong evidence; effective evaluation
	Reasoning	Basic logical links	Logical and mostly consistent	Strong, persuasive reasoning
	Expression modes	Limited use of modes	Uses two or more modes appropriately	Integrates multiple modes effectively
	Language use	Frequent errors (6–10)	Minor errors (3–5)	Minimal errors (less than 3); precise language
Organization	Structure completeness	Incomplete or imbalanced	Complete with minor issues	Clear introduction, body, and conclusion; well-balanced
	Coherence	Limited coherence	Generally coherent	Strong coherence with effective transitions
Creativity	Ideas	Basic or conventional ideas	Some original insight	Multiple original, insightful ideas
	Writing techniques	Minimal creativity	Some creative expression	Highly effective and original style

2. Methodology

The research team conducted a pilot study to evaluate the feasibility of the proposed writing competency assessment standards for Grade 10 students in real-world conditions. The trial was conducted with Grade 10 students from two upper secondary schools in Hải Dương Province: one located in a rural area (Quang Trung Upper Secondary School, Ninh Giang District) and one in an urban area (Hồng Quang Upper Secondary School, Hải Dương City). A total of 164 students participated in the study, with 89 from the rural school and 75 from the urban school. This study was conducted as a pilot validation. Although the sample was limited to two schools (n = 164), it was sufficient to examine feasibility, internal consistency, and structural coherence of the proposed assessment standards rather than national generalizability.

To evaluate the feasibility of the proposed writing competency assessment standards, we conducted a pilot test with Grade 10 students using an open-ended essay question: "Suppose you have a close friend who frequently stays up late until 1–2 a.m. Write an argumentative essay of approximately 800–1000 words to persuade your friend to break this habit."

Based on theoretical research, writing competency at the upper secondary level includes four key criteria:

- Clearly expressing the main idea.
- Developing and presenting content appropriate to the argumentative genre.
- Ensuring proper essay structure.
- Demonstrating creativity in writing.

Each of the four main criteria is further subdivided into a total of twelve sub-criteria. Each sub-criterion is evaluated using a five-point scale, with codes ranging from 0 to 4. These codes correspond to specific score values, such that code 0 represents 0 points, code 1 corresponds to 0.25 points, code 2 represents 0.5 points, code 3 corresponds to 0.75 points, and code 4 represents the maximum value of 1.0 point.

The 12 sub-criteria were analyzed and evaluated as a measurement scale, serving as the basis for analysis using the following methods:

- Internal reliability was examined using Cronbach's Alpha coefficient.
- Structural reliability was assessed through Exploratory Factor Analysis (EFA) and the Rasch model.
- Test scores were statistically analyzed and described using both a 10-point grading scale and a competency-based score scale.
- Student competency levels were evaluated using Guttman scaling software.

The twelve sub-criteria, their corresponding main criterion codes, and their operational descriptions are presented in Table 2.

Table 2: Coding of sub-criteria for assessing writing competency in argumentative essays

No.	Criterion code	Main criterion	Sub-criterion code	Sub-criterion description
1	TC1	1. Clearly expressing the main idea	C1.1	Accurately identifies the argumentative issue and focuses on the core content of the essay.
2			C1.2	Ensures accuracy of content knowledge and the information.
3			C1.3	Expresses appropriate viewpoints that reflect the writer's understanding of the issue.
4	TC2	2. Developing and presenting content appropriate to the argumentative genre	C2.1	Constructs a logical and coherent argument system.
5			C2.2	Uses representative reasoning and evidence to support the argument system.
6			C2.3	Constructs coherent arguments using appropriate reasoning techniques.
7			C2.4	Appropriately and effectively uses modes of expression.
8			C2.5	Meets language and formatting requirements in expression, word choice, sentence construction, handwriting, and presentation.
9	TC3	3. Ensuring proper essay structure	C3.1	Presents a complete structure with appropriate length for each section.
10			C3.2	Maintains coherence and logical connections across sections and paragraphs.
11	TC4	4. Creativity in writing	C4.1	Demonstrates creativity in content/ideas.
12			C4.2	Demonstrates creativity in writing techniques (word choice, sentence structure, rhetorical devices, formatting, etc.).

After the data collection, students' responses were analyzed using SPSS version 22.0. The internal reliability of the scale was examined using Cronbach's Alpha coefficient. The results are presented in Table 3.

According to the SPSS analysis results (Table 3), the reliability coefficients for the subscales corresponding to each criterion are relatively high. The criteria of clearly expressing the main idea, developing and presenting content appropriate to the genre, and demonstrating creativity in writing show good reliability, as indicated by Cronbach's Alpha values exceeding 0.80. However, the reliability coefficient for the criterion, Ensuring proper essay structure, is lower (Cronbach's Alpha = 0.68), possibly due to only two sub-criteria or the lack of strong internal consistency between them. This may indicate that the sub-criteria differ in how they reflect the construct or that the indicators themselves are not sufficiently well-defined, leading to inconsistencies in student performance across items. The lower reliability coefficient for the essay structure criterion (α = 0.68) is likely due to the small number of sub-criteria (two items). For a pilot instrument, this level is acceptable, and the indicators were retained for their pedagogical relevance. The overall reliability coefficient of the writing competency assessment scale is high (Cronbach's Alpha = 0.92), indicating strong internal consistency among the 12 sub-criteria and confirming that they measure the same construct - writing competency. Additionally, the analysis of

how each sub-criterion relates to the total score and the "alpha if item deleted" checks shows that all sub-criteria are essential for the scale. None of the sub-criteria were flagged for removal, confirming their relevance and appropriateness within the overall measurement framework. In summary, the reliability analysis demonstrates that all sub-criteria contribute significantly to the overall scale and are closely correlated with the total score in measuring a familiar construct - students' writing competency. The scale can be considered a reliable tool for assessing student writing ability, and the use of its results in further analyses is appropriate and dependable. EFA was performed using SPSS version 22.0, employing the Principal Components method, to assess the reliability of the scale in measuring students' writing skills. The primary objective of the factor analysis was to examine whether the sub-criteria measured one-dimensionality or multidimensionality. In other words, the purpose of factor analysis is to identify the latent structure of the scale by detecting common factors that account for the correlations among the observed variables, thereby eliminating the influence of unrelated factors. The results of the factor analysis based on student response data are shown in Table 4.

Table 3: Reliability analysis of the scale

Criterion/subscale	Cronbach's alpha	Number of sub-criteria
1. Clearly expressing the main idea	0.89	3
2. Developing and presenting content appropriate to the genre	0.88	5
3. Ensuring proper essay structure	0.68	2
4. Creativity in writing	0.83	2
Overall scale	0.92	12

Table 4: KMO measure and Bartlett's test of sphericity

KMO coefficient	Bartlett's test of sphericity
0.787	0.000

Suppose the KMO coefficient is greater than 0.5 and Bartlett's Test of Sphericity yields a significant result ($p < 0.05$). In that case, it indicates that the items are well correlated in the population, confirming that factor analysis can be applied. According to Table 4, the conditions for analyzing the questionnaire data are met. Furthermore, the results of the factor analysis show that the total variance explained by the extracted components is 64.91%, based on the 12 items of the scale. Since this value exceeds the acceptable threshold of 50%, it confirms that the scale demonstrates strong explanatory power. Regarding the number of principal components extracted from the 12 sub-criteria of the scale: In this study, factor analysis was performed using the Principal Components method with Varimax rotation, applied to all 12 sub-criteria. The results of the factor extraction are illustrated in Fig. 1. According to theoretical guidelines, if the eigenvalue of the first factor is approximately four times greater than that of the second, the items in the scale can be considered to measure a single common factor. The analysis results show a clear gap between the first and second factors, with the first factor's eigenvalue nearly four times that of the second. This supports the conclusion that most of the sub-criteria contribute to representing a single latent factor and that the scale's components collectively reflect a common factor.

3. Results

3.1. Rasch model analysis

The writing assessment scale for competency is polytomous. Therefore, the appropriate model for analyzing the pilot test data is the Partial Credit Model (PCM). In this study, we utilized the Quest software, developed by ACER, which is based on Item Response Theory (IRT) and the generalized multidimensional Rasch Item Response Model (Rasch, 1960).

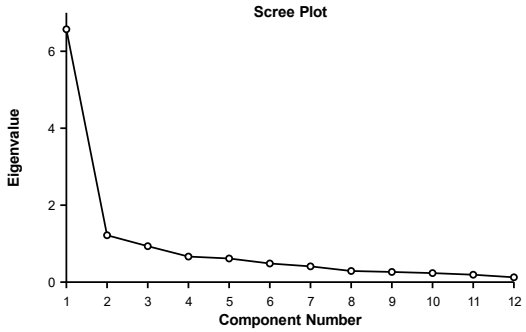


Fig. 1: Component number graph

Using IRT estimation, students' answers on the written test were analyzed in Quest to look at several factors: how well each part fits the Rasch model, how students' skill levels relate to the difficulty of each part, how reliable the assessment scale is, and essential item statistics like item difficulty, discrimination index, and how each part relates to the overall test score. The results from the IRT analysis provide evidence on how the proposed assessment standards represent writing competency and how well each subcriterion aligns with student performance levels. To assess the reliability of the writing competency scale, we employed the Rasch model from IRT theory and found that the 12 sub-criteria formed a cohesive group, with a confidence interval of 0.77–1.33 (Fig. 2). Based on the quantitative analyses of students' test performance, the set of writing competency assessment criteria demonstrates strong reliability. Each sub-criterion is closely related and plays an essential role within the overall measurement scale. The general structure—comprising four main criteria and twelve sub-criteria—was designed to measure a single construct consistently. This confirms the scale's appropriateness for assessing students' writing competency, as intended.

3.2. Descriptive statistics

Table 5 presents the descriptive statistics based on the 10-point scale for the total score of all 12 assessment criteria.

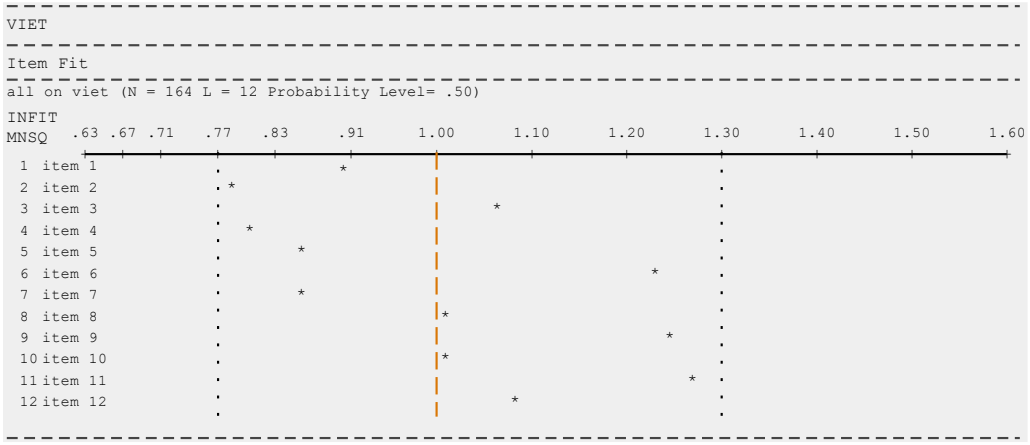


Fig. 2: Item fit statistics for the Rasch model

Table 5: Descriptive statistics and score distribution of the test (N = 164)

Statistic	Value
Mean score	5.58
Median score	5.75
Mode	5.50
Standard deviation	1.74
Skewness	-0.65
Standard error of skewness	0.19
Kurtosis	-0.40
Standard error of kurtosis	0.38
Minimum score	0.80
Maximum score	9.50

The students' average test score was 5.58 out of 10, indicating a moderate to good performance level. The most frequently obtained score (mode) was 5.50, the median was 5.75, while the lowest and highest scores were 0.80 and 9.50, respectively. The skewness value of -0.65 suggests a slightly left-skewed distribution, meaning that more students achieved scores above the mean than below it. The kurtosis value of -0.40 indicates a flatter distribution compared to the normal curve, suggesting the scores were more evenly spread rather than heavily concentrated around the average.

Overall, the students' test scores show a relatively good distribution, broad and close to a normal curve, reflecting the actual range of student writing competency. The descriptive statistics provide initial evidence that the pilot test results reasonably reflect students' writing ability. More importantly, the results suggest that the test was appropriate for both student groups from rural and urban school settings.

3.3. Descriptive statistics by individual criteria

As previously discussed, writing competency consists of four main criteria. These criteria have been confirmed to demonstrate strong reliability and are part of a unified structural construct. The average scores for the four criteria, calculated on a 10-point scale, are presented in Table 6.

Table 6: Descriptive statistics for the four writing competency criteria

Statistic	Criterion 1	Criterion 2	Criterion 3	Criterion 4
Mean	6.32	5.35	3.80	1.27
Median	6.67	6.00	3.75	1.25
Mode	5.00	6.00	3.75	1.25
Standard deviation	2.20	1.64	1.45	1.06
Minimum score	1.67	0.00	0.00	0.00
Maximum score	10.00	8.50	6.25	5.00

According to the 10-point scale, the scores of the four writing competency criteria show clear variation. The average score decreases progressively from Criterion 1 (6.32) to Criterion 4 (1.27), with both the median and mode scores reflecting the same trend. Criterion 1 also has the highest standard deviation (2.20), indicating greater score dispersion compared to the other criteria. In contrast, Criterion 4 shows more concentrated scores, with a lower standard deviation of 1.06. The minimum score for the last three criteria was 0.00, suggesting that some students did not complete or did not meet the requirements of these criteria. Meanwhile, the maximum scores also decreased from 10.00 (Criterion 1) to 5.00 (Criterion 4).

Criteria 1 and 2 had relatively high average scores (6.32 and 5.35, respectively), while Criteria 3 and 4 recorded notably lower averages (3.80 and 1.27). These differences suggest the need to review the content and expectations of Criteria 3 and 4 to understand why student performance was weaker. Additionally, it may be necessary to examine teaching and learning practices in schools related to the specific skills assessed by these two criteria.

To examine the appropriateness of the proposed assessment standards, we conducted a pilot study at two schools representing different regions: one rural school (School 1 – coded as Rural School, RS) and one urban school (School 2 – coded as Urban School, US). The results indicate a significant difference in overall mean scores between rural and urban students.

The differences in mean scores across the four assessment criteria were analyzed using an independent-samples t-test, and the results are presented in Table 7.

Table 7: Differences in mean scores of the four criteria by schools

Criteria	School	N	Mean	Standard deviation	Mean difference	T	p
1. Clearly expressing the main idea	RS	89	5.58	2.28	-1.62	-5.05	0.00
	US	75	7.20	1.73			
2. Developing and presenting content appropriately	RS	89	4.92	1.66	-0.93	-3.78	0.00
	US	75	5.85	1.47			
3. Ensuring proper essay structure	RS	89	3.74	1.69	-0.15	-0.65	0.52
	US	75	3.88	1.10			
4. Creativity in writing	RS	89	1.17	0.84	-0.22	-1.31	0.19
	US	75	1.38	1.28			

The results indicate a significant difference in overall writing competency between students from rural and urban schools, with an average score gap of 1.02 points ($p < .05$). When analyzed by individual criteria, significant differences were found in Criterion 1 (Clearly expressing the main idea) and Criterion 2 (Developing and presenting content appropriately), both at the $p < .05$ level. However, for Criterion 3 (Ensuring proper essay structure) and Criterion 4 (Creativity in writing), the results did not provide sufficient evidence to confirm differences between the two school groups. These findings reflect differentiation across the assessment criteria.

Socioeconomic conditions and educational disparities may explain why rural students perform worse than their urban counterparts in meeting the proposed writing competency standards. Nevertheless, the observed gap is not large enough to compromise the applicability of the assessment standards. This suggests that the proposed writing competency framework remains inclusive and feasible for both rural and urban student populations.

3.4. Correlation between student ability and item difficulty based on IRT

Using Item Response Theory (IRT) analysis with QUEST software, Fig. 3 presents a comparison between students' abilities (indicated by "X" on the left) and the difficulty levels of the assessment items, both displayed on a scale ranging from -4 to 4. Students whose ability level is lower than, equal to, or higher than an item's difficulty level will have a lower than 50%, approximately 50%, or greater than 50% probability of successfully completing that item, respectively.

Fig. 3 illustrates that with an average item difficulty of 0, the majority of students fall within an ability range of approximately -5 to 5 logits. The distribution of item difficulty (ranging from -6 to 6) is relatively well-aligned with student ability levels. Items (i.e., sub-criteria) are ordered from easiest to most difficult, providing an appropriate spread for competency-based assessment. The rating levels associated with the 12 sub-criteria effectively cover a wide range of abilities, with most values spanning from -5 to 5. This suggests that the full range of student abilities is adequately represented, and no ability levels are left unassessed by the current set of items.

An analysis of item discrimination and difficulty across the four scoring levels (0–4) reveals high discrimination across all 12 sub-criteria, with classical discrimination indices all exceeding 0.5. The difficulty levels of each rating step are appropriate:

- Lower scores (levels 0 and 1) were achieved by fewer students and were associated with negative point-biserial correlations (Pt-bis < 0),
- Higher scores (especially level 2) were also rare but showed positive Pt-bis values (Pt-bis > 0).

This indicates that higher scores correctly reflect higher levels of student ability. Students with higher writing competency are more likely to achieve level 2, while those with lower ability tend to remain at levels 0 or 1.

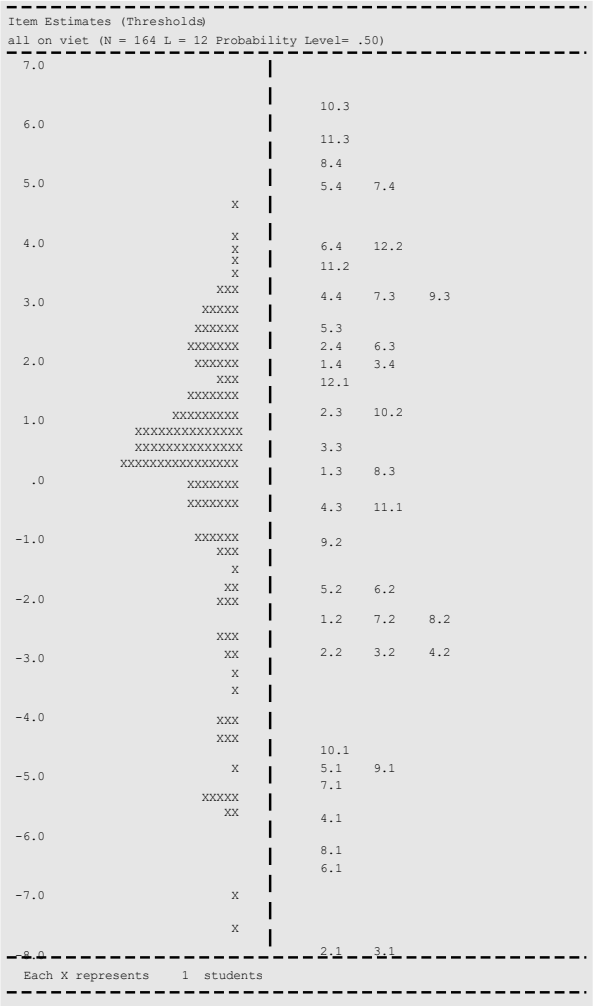


Fig. 3: Correlation between student ability and item difficulty

Based on these findings, it can be concluded that the criteria for assessing writing competency are scientifically sound and well-constructed. The 12 sub-criteria demonstrated high reliability

(Cronbach's Alpha > 0.90), indicating strong internal consistency. The EFA results further confirmed a coherent and valid factor structure aligned with the intended assessment goals. Additionally, the Quest-based IRT analysis showed that all sub-criteria fit together into one main idea, clearly distinguished between different student groups, and matched the overall range of student abilities.

In summary, the writing competency assessment criteria satisfy the key requirements of reliability, validity, and structural coherence, and are well-suited for evaluating student writing performance in upper secondary education.

3.5. Proficiency levels in argumentative writing competency

Based on the results from the Guttman (1944) scale analysis, students were categorized into four proficiency levels: excellent, good, average, and below average (Fig. 4). The sub-criteria were ranked in descending order of achievement, from highest (C1.1) to lowest satisfactory. Due to space limitations in this article, we present a simplified illustration by selecting a sample of 20 students, with five students representing each proficiency level, to demonstrate how performance is distributed across the Guttman scale.

The group of students classified as excellent (score ≥ 8.0) included 12 out of 164 students, accounting for 7.31% of the sample. Among these students, the highest-performing sub-criteria were C1.1 (accurately identifying the argumentative issue and maintaining focus on the main idea) and C1.3 (expressing a clear and appropriate viewpoint that reflects understanding of the issue). The lowest scores were observed in C4.1 (creativity in content and ideas) and C4.2 (creativity in writing skills such as expression, word choice, sentence structure, and use of rhetorical devices), with three students receiving a score of 0 on both.

STT	C1.1	C1.3	C2.1	C1.2	C2.5	C2.3	C2.4	C2.2	C3.1	C3.2	C4.1	C4.2	Tổng
1	1	1	0.75	1	0.75	0.75	0.75	0.75	0.5	0.5	0.5	0	8.25
2	0.75	1	0.75	0.75	0.75	1	0.75	0.75	0.5	0.5	0.25	0.5	8.25
3	0.75	1	0.75	0.75	0.75	1	0.75	0.75	0.5	0.5	0.25	0.5	8.25
4	1	1	0.75	0.75	1	0.75	0.75	0.75	0.5	0.5	0.25	0	8
5	1	1	0.75	0.75	0.75	1	0.75	0.75	0.5	0.25	0.5	0	8
6	1	0.5	0.75	1	0.75	0.5	0.5	0.5	0.5	0.5	0.25	0.5	7.25
7	0.75	1	0.75	0.75	0.75	0.5	0.75	0.75	0.5	0.25	0.25	0	7
8	0.75	1	0.75	1	0.75	0.5	0.5	0.5	0.5	0.5	0	0.25	7
9	0.75	1	0.75	0.75	0.75	0.5	0.5	0.5	0.5	0.5	0.25	0.25	7
10	0.75	1	0.75	0.5	0.5	0.75	0.75	0.75	0.5	0.25	0.25	0	6.75
11	1	0.5	0.75	0.5	0.75	0.5	0.5	0.5	0.5	0.25	0	0.25	6
12	1	0.5	0.75	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.25	0	6
13	1	0.5	0.75	0.75	0.75	0.25	0.5	0.5	0.5	0.25	0	0.25	6
14	0.75	0.75	0.5	0.75	0.5	0.75	0.5	0.5	0.5	0.25	0	0	5.75
15	0.75	0.75	0.5	0.5	0.5	0.75	0.5	0.75	0.5	0.25	0	0	5.75
16	0.5	0.75	0.5	0.5	0.5	0.5	0.5	0.5	0.25	0.25	0	0	4.75
17	0.75	0.75	0.5	0.75	0.25	0.5	0.5	0.25	0.25	0.25	0	0	4.75
18	0.5	0.25	0.5	0.5	0.75	0.25	0.5	0.25	0.5	0.5	0	0.25	4.75
19	0.5	0.5	0.75	0.5	0.5	0.5	0.5	0.5	0.25	0.25	0	0	4.75
20	0.5	0.5	0.5	0.5	0.75	0.5	0.25	0.25	0.5	0.25	0.25	0	4.75

Fig. 4: Guttman scale analysis

The good group (scores from 6.5 to 7.9) included 36 students (21.95%). The score distributions across most sub-criteria suggest a relatively balanced writing ability among students in this group. However, on sub-criteria C2.3 (construct sub-criteria arguments and using reasoning techniques) and C2.4 (effectively using appropriate methods for sub-criteria C2.3 (making clear arguments and using reasoning) and C2.4 (using the right ways to express ideas), we still face challenges with these specific competencies. According to Vygotsky (1978), this reflects their "zone of proximal development," highlighting the need for targeted pedagogical support to help students meet these standards. Notably, more than two-thirds of the group did not achieve either of the two creativity-related sub-competencies.

The average group (scores from 5.0 to 6.4) accounted for 78 students, or 47.56% of the total sample. Their results suggest relative strength in sub-criteria such as identifying the main issue, maintaining focus, expressing a sound viewpoint, and outlining appropriate sub-criteria. However, consistent with broader trends, very few students in this group demonstrated creativity in writing skills, with most scoring below the average level (0.25 out of 0.5).

The below-average group (scores below 5.0) included 38 students, making up 23.17% of the sample. Most sub-criteria in this group fell below standard, with the lowest performance again found in creativity-related competencies. Only one student in this group reached the average level (0.25) on C4.2, specifically in the creativity of writing technique. This implies that nearly all other students in this group achieved the average level (0.25) for C4.2, which assesses the attempt to apply knowledge of the Vietnamese language (e.g., experimenting with word choice, sentence structure, or rhetorical devices). Further examination of their writing revealed common issues such as: failing to focus on the argumentative goal (e.g., persuading a friend to change behavior rather than simply discussing a bad habit); incomplete or unclear argument structures;

poor organization; weak coherence between paragraphs and sentences; insufficient use of supporting evidence; lack of appropriate expressive techniques; mismatched viewpoints or examples; and careless presentation (e.g., illegible handwriting, excessive erasures). Low performance in creativity-related criteria likely reflects exam-oriented writing instruction that prioritizes structure and content accuracy over originality. The standards therefore function diagnostically by identifying creativity as an underdeveloped component of writing competency.

The Guttman scale distribution also indicates that students with identical total scores may exhibit different performance profiles across individual competency sub-criteria. For example, three students who each obtained a total score of 8.25 demonstrated different patterns of achievement. Student 1 achieved the maximum score for C4.1 (creativity in content and ideas), demonstrating an ability to present personal and insightful perspectives. However, this student received no points for C4.2 (creativity in writing techniques), suggesting that creativity was not equally demonstrated across the two dimensions. The student also achieved the maximum score for C1.2 (accuracy of content knowledge). In contrast, Students 2 and 3 achieved the maximum scores in both C4.1 and C4.2, indicating stronger performance across the two creativity-related sub-criteria. They also achieved higher scores in C2.3 (constructing coherent arguments using appropriate reasoning techniques) than Student 1. These examples illustrate that identical overall scores can conceal differences in students' competency profiles. Consequently, examining performance at the sub-criterion level can provide more detailed diagnostic information than relying exclusively on total scores.

These differences demonstrate that even students with identical total scores require individualized instruction and support. By analyzing performance at the sub-criterion level, educators can tailor learning strategies to help each student build a more comprehensive and balanced writing competency.

4. Discussion

The Rasch analysis indicates that item difficulty was appropriately aligned with student ability levels, supporting the developmental interpretation of the standards. Higher-level criteria captured more complex writing competencies, while lower-level criteria reflected foundational skills. We found that the proposed writing competency standards for upper secondary students are feasible. Most students met the baseline level (level 1) across the criteria, and some students demonstrated higher levels of proficiency (level 2) in multiple sub-criteria. The standards also proved to be broadly applicable to students at both participating schools, which differ in terms of geographic location, socioeconomic conditions, infrastructure, and teaching resources. In addition, the standards supported differentiation, helping to identify strengths and weaknesses across student groups based on the defined components and performance levels of writing competency.

To improve the reliability of the sub-criteria in the proposed writing assessment standards, further attention should be given to the actual difficulty of specific sub-criteria. Sub-criteria may be overlapping or too challenging for students, for example, criteria 1 and 3, or 11 and 12. It may also be necessary to develop additional tools to assess writing competency at different time points, enabling longitudinal comparisons and more detailed identification of student progress—something that could not be carried out within the scope of this article.

The results show that students performed relatively well in three core criteria: clearly expressing the main idea, developing and presenting content appropriate to the essay type, and ensuring proper essay structure. This indicates that students have acquired essential argumentative writing skills. However, the criterion of creativity in writing was not achieved by most students. This highlights potential directions for the future. However, most students failed to meet the criterion of creativity in writing. Documentary writing in particular, in Vietnamese secondary schools. It also suggests that assessment instruments (e.g., test prompts) should be designed in ways that better stimulate and evaluate students' creative capacity when engaging with given tasks.

With the diagnostic data visualized on the Guttman scale, teachers can identify students' competency levels for each sub-criterion, compare different groups of students across the development spectrum, and analyze students at the same overall score level but with varying patterns of achievement. This allows educators to locate both the current level of performance and the "zone of proximal development" for each learner, thereby informing targeted pedagogical interventions and guidance. Such insight helps teachers support students in reinforcing their strengths and addressing areas for improvement in their writing competency in subsequent learning phases.

Following the trial, we conducted a review and made adjustments to certain aspects of the standards. Specifically, we revised the wording of some performance levels to make them more explicit and better aligned with observable indicators of student performance. We also refined the scope of criteria 1 and 3 for clarity.

To scale the competency standards for argumentative essays nationally, further research is needed to conduct broader trials across all upper secondary grades (10, 11, and 12). These trials should be expanded across different regions to ensure more comprehensive data, allowing for more informed adjustments. We should expand these trials across various areas to gather more comprehensive data that can guide appropriate adjustments to the standards. Writing standards are based on the expected learning outcomes of the

General Education Curriculum 2018. This line of research will enhance the overall effectiveness of language assessment in Vietnamese secondary schools as the 2018 curriculum continues to be implemented.

The proposed writing standards can also serve as a reference point for teachers to communicate their expectations to students regarding argumentative writing proficiency. Additionally, the standards offer a basis for teachers to locate a student's current level of writing ability and design instructional plans that help students progress toward higher levels. For students, the standards provide a framework for self-assessment after completing a written task, enabling them to reflect on which component skills require further development. In this way, the standards function as a diagnostic tool, helping to make assessment more objective, transparent, and pedagogically meaningful in upper secondary schools. This study focused on quantitative validation to establish psychometric reliability. Qualitative investigation of classroom implementation will be addressed in future research.

5. Conclusion

The development of writing competency assessment standards in Vietnamese Literature is a necessary task. It serves as a pedagogical guide for teachers in organizing instruction and assessment within the general education system, aligned with the goals of developing student competencies and qualities as outlined in the General Education Curriculum 2018 in Vietnam. The process of building these standards should follow a systematic procedure - from defining the target competency and its components, behavioral indicators, and performance levels to piloting the proposed standards.

Once established, the writing competency assessment standards will provide teachers and schools with a concrete foundation for designing instruction and evaluation with the curriculum's requirements. Moreover, such standards will allow teachers to offer timely intervention and support, helping students progress through successive levels of competency development. This reflects the principle of teaching and assessment for student growth. However, at present, the development of competency-based assessment standards for upper secondary students in Vietnamese Literature, particularly for writing competency, remains largely at the research and experimental stage. In practice, assessing writing competency continues to pose significant challenges for teachers. Therefore, educational authorities should take active steps to accelerate the process of making writing competency a focus.

This is a critical step in implementing the reformed curriculum and will empower both teachers and students to monitor better and enhance the effectiveness of teaching and learning.

List of abbreviations

ACER	Australian Council for Educational Research
EFA	Exploratory factor analysis
INFIT	Information-weighted fit statistic
IRT	Item response theory
KMO	Kaiser-Meyer-Olkin measure of sampling adequacy
L	Number of items (12 sub-criteria)
MNSQ	Mean-square statistic
MOET	Ministry of Education and Training of Vietnam
N	Number of students (sample size)
PCM	Partial credit model
Pt-bis	Point-biserial correlation
RS	Rural school
SPSS	Statistical Package for the Social Sciences
STT	Serial number
US	Urban school
ZPD	Zone of proximal development

Compliance with ethical standards

Ethical considerations

The study was conducted in accordance with relevant ethical guidelines. Permission was obtained from the participating schools, and informed consent was secured from participants and their parents or legal guardians, where applicable. Students' confidentiality and anonymity were maintained throughout the study.

Conflict of interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

Delors J (1996). Treasure within: Highlights. UNESCO Publishing, Paris, France.

Guttman L (1944). A basis for scaling qualitative data. American Sociological Review, 9(2): 139-150. https://doi.org/10.2307/2086306

Hope D, Kluth D, Homer M, Dewar A, Goddard-Fuller R, Jaap A, and Cameron H (2025). Exploring the use of Rasch modelling in "common content" items for multi-site and multi-year assessment. Advances in Health Sciences Education, 30(2): 427-438. https://doi.org/10.1007/s10459-024-10354-y **PMid:38977526 PMCid:PMC11965148**

Imsa-ard P (2025). Learning-oriented assessment and L2 argumentative writing: Effects on writing ability and academic resilience. Discover Education, 4: 396. https://doi.org/10.1007/s44217-025-00779-x

Mojumdar P (2026). The critical extension framework: Making organisational assumptions explicit in multilingual writing assessment. Language Testing in Asia, 16: 42. https://doi.org/10.1186/s40468-026-00458-0

Ngo HTD, Nguyen QT, and Smith RJ (2024). Literacy education in Vietnamese schooling system. In: Ha PL, Bao D, and Windle J (Eds.), Vietnamese language, education and change in and outside Vietnam: 13–31. Springer, Singapore, Singapore. https://doi.org/10.1007/978-981-99-9093-1_2

Phuong HY, Phan QT, and Le TT (2023). The effects of using analytical rubrics in peer and self-assessment on EFL students' writing proficiency: A Vietnamese contextual study. Language Testing in Asia, 13: 42. https://doi.org/10.1186/s40468-023-00256-y

Poehner ME (2025). In pursuit of pink elephants: Sociocultural Theory and the determination of learning in assessments that support L2 teaching and learning. System, 129: 103585. https://doi.org/10.1016/j.system.2024.103585

Ponomariovienė J, Jakavonytė-Staškuvienė D, and Torterat F (2025). Implementing competency-based education through the personalized monitoring of primary students' progress and assessment. Education Sciences, 15(2): 252. https://doi.org/10.3390/educsci15020252

Rasch G (1960). Probabilistic models for some intelligence and attainment tests. Danmarks Pædagogiske Institut, Copenhagen, Denmark.

Tang C, Xiong J, and Engelhard G (2025). Identification of writing strategies in educational assessments with an unsupervised learning measurement framework. Education Sciences, 15(7): 912. https://doi.org/10.3390/educsci15070912

Vygotsky LS (1978). Mind in society: The development of higher psychological processes. Harvard University Press, Cambridge, USA.

Yen PH and Thao LT (2024). Exploring the implementation and perception of competency-based assessment practices among Vietnamese EFL instructors. Language Testing in Asia, 14(1): 26. https://doi.org/10.1186/s40468-024-00300-5