QUBAHAN ACADEMIC JOURNAL
VOL. 6, NO. 3, July 2026
https://doi.org/10.48161/qaj.v6n3a2444

Reconceptualizing Critical Thinking in Academic Writing: A Rasch Model Investigation of Cognitive, Linguistic, and Cultural Dimensions in Higher Education

Khabib Sholeh [1], Onok Yayang Pamungkas [2*], Herlinawati [3], Marhanani Tri Astuti [4], Ahmad Dudin [5] and Aleeyah Masae [6]

[1] Department of Indonesian Language and Literature Education, Faculty of Teacher Training and Education, Universitas Muhammadiyah Purworejo, Purworejo 54111, Indonesia;
[2] Department of Indonesian Language and Literature Education, Faculty of Teacher Training and Education, Universitas Muhammadiyah Purwokerto, Purwokerto 53182, Indonesia;
[3] Research Center for Language, Literature, and Community, National Research and Innovation Agency (BRIN), Jakarta, 12710, Indonesia;
[4] Research Centre for Industrial Economics, Services, and Trade, National Research and Innovation Agency (BRIN), Jakarta, 12710, Indonesia;
[5] Research Center for Educational, National Research and Innovation Agency (BRIN), Jakarta, 12710, Indonesia; email:
[6] Malay Language Education Study Program, Faculty of Humanities and Social Sciences, Yala Rajabhat University, Yala, 95000, Thailand.

* **Corresponding author:** onokyayangpamungkas@ump.ac.id.

ABSTRACT: This study aims to (1) validate a measurement instrument for assessing critical thinking in academic writing using the Rasch model and (2) develop a theoretical understanding of critical thinking as a multidimensional construct shaped by the interaction between cognitive reasoning, linguistic mediation, and cultural interpretation. The study involved 257 undergraduate students in an Indonesian higher education context, using a 36-item Likert-scale instrument. Rasch analysis was employed to evaluate measurement reliability, item functioning, dimensionality, and person–item targeting. The results indicate moderate empirical reliability alongside stronger model-based reliability, suggesting stable latent estimation within the probabilistic framework. Item fit analysis revealed several misfitting items, while residual Principal Component Analysis (PCA) indicated the presence of secondary dimensions beyond the primary construct. These findings support the interpretation that critical thinking in academic writing is not strictly unidimensional. Theoretically, this study contributes by proposing an integrative measurement perspective that reconceptualizes critical thinking as a construct emerging from the interaction of cognitive, linguistic, and sociocultural dimensions within academic discourse. Pedagogically, the findings highlight the need for instructional approaches that integrate language proficiency and cultural awareness in developing critical thinking in writing. However, the findings are limited to the specific sample and context studied, and further validation across diverse populations is required.

Keywords: Language–culture interface, Critical thinking, Academic discourse, Rasch measurement model, Higher education, Multidimensional construct.

I. INTRODUCTION

The development of global higher education places critical thinking skills as an essential competency in academic writing, especially in the context of increasingly complex and cross-cultural knowledge production. Academic writing is no longer just an activity of reproducing information, but a process of constructing meaning that involves the integration of linguistic ability, argumentative structure, and sensitivity to socio-cultural contexts [1-4]. Within this framework, language functions not only as a medium of communication, but also as an epistemological tool that shapes ways of thinking [5, 6]. On the other hand, culture provides an interpretive framework that influences the way individuals understand, organize, and evaluate information [7, 8]. However, learning practices in higher education still tend to separate the dimensions of language, culture, and cognition, so that the development of critical thinking in academic writing is often partial and not integrated.

Although a number of studies have examined the relationship between critical thinking and academic writing, most still view critical thinking as a single cognitive construct separate from linguistic and cultural factors [9-11]. Recent empirical studies show that critical thinking performance in writing is strongly influenced by language skills, metacognitive strategies, and the socio-cultural context behind which the author [12-14]. However, there are methodological gaps in the measurement of these constructs, especially regarding the assumption that assessment instruments measure a single underlying construct. Traditional measurement approaches tend to ignore the complexity of latent structures that are multidimensional, thus potentially resulting in less accurate estimates [15-17]. In addition, research that integrates Rasch's model-based psychometric analysis with language and cultural perspectives in the context of academic writing is still relatively limited, especially in the context of Indonesian higher education.

To address these methodological limitations and to capture the complexity of the construct under investigation, this study adopts the Rasch measurement model as a theoretically grounded analytical framework. Conventional approaches, such as Classical Test Theory (CTT), which rely primarily on observed test scores, have limited capacity to provide invariant measurement and to examine the interaction between respondent ability and item difficulty at the latent level. In contrast, the Rasch model estimates person ability and item difficulty on a common logit scale, thereby supporting parameter invariance and comparability across individuals and items [18]. More importantly, the Rasch model provides detailed diagnostic tools, including item fit, person fit, rating scale functioning, and residual-based Principal Component Analysis (PCA), which can be used to examine whether additional latent dimensions remain after the primary measurement dimension has been modeled. Although the Rasch model is commonly associated with the assumption that an instrument measures a single underlying construct, recent developments in multidimensional Rasch modeling demonstrate its capacity to examine complex latent structures in educational measurement, particularly when constructs involve interacting cognitive and contextual dimensions [19, 20]. Given that this study aims not only to validate an instrument but also to investigate whether critical thinking in academic writing reflects a multidimensional construct shaped by linguistic and cultural factors, the Rasch approach offers a robust framework to detect misfit patterns and latent structural complexity. Therefore, the use of the Rasch model is not merely a technical choice, but a theoretically aligned methodological strategy to address the complexity of the research problem. However, methodological gaps remain in the measurement of these constructs, particularly because assessment instruments often assume that the measured ability reflects only one underlying construct.

Based on these gaps, this research offers conceptual and methodological novelty by examining critical thinking skills in academic writing as a multidimensional construct formed through the interaction between language, culture, and cognitive processes. Methodologically, this study uses the Rasch model approach to evaluate the measurement quality, item function, dimensionality structure, and suitability between respondents and instruments. Conceptually, this study proposes an integrative model that positions the interaction between language and culture as the foundation in the formation of critical thinking patterns. Thus, the objectives of this study are to (1) evaluate the quality of critical thinking measurement instruments in academic writing, (2) identify the underlying construct structure, and (3) develop a theoretical understanding of the relationship between language, culture, and critical thinking in the context of higher

education. Ultimately, this research contributes by integrating Rasch-based psychometric analysis with a sociocultural perspective in academic writing, which has been rarely done in previous studies.

II. THEORETICAL FRAMEWORK

1. CRITICAL THINKING IN ACADEMIC WRITING

Critical thinking in academic writing is a complex cognitive process that involves the ability to analyze, evaluate, and synthesize information to produce logical and evidence-based arguments [21-24]. In the context of higher education, this ability is not only related to higher-level thinking skills, but also to the ability to systematically articulate ideas through academic language [25, 27]. Academic writing, thus, becomes an arena in which critical thinking processes are materialized in the form of text structure, argumentation, and linguistic choices.

Nevertheless, recent research shows that critical thinking in writing is not universal or homogeneous, but rather is influenced by a variety of factors, including language skills, metacognitive strategies, as well as the cultural context of the writer [28-30]. This confirms that critical thinking in academic writing is not only a cognitive phenomenon, but also a linguistic and social phenomenon. Therefore, an approach that views critical thinking as a single construct becomes inadequate in explaining the complexity of academic writing practices in multicultural settings.

2. LANGUAGE AS A MEDIUM OF COGNITION AND REPRESENTATION OF KNOWLEDGE

In a sociocultural perspective, language serves as a primary mediating tool in the formation and development of thought processes [31, 32] Language not only represents thoughts, but also shapes the cognitive structure of individuals through social interaction. Correspondingly, the systemic functional linguistic approach places language as a semiotic system that realizes meaning in a given social context [33-35]. In academic writing, lexical choices, syntactic structure, and discourse organization reflect and shape the writer's thought process.

Empirical research shows that the level of language proficiency has a significant influence on the quality of reasoning in academic writing [28, 36-38]. In addition, limitations in linguistic competence can lead to distortions in the expression of ideas, which ultimately affects critical thinking performance [39, 40]. Thus, language is not just a means of communication, but an integral component in the construction of critical thinking. This perspective is important in understanding why variations in language ability can result in variations in the quality of academic argumentation.

3. CULTURE AS AN INTERPRETIVE FRAMEWORK IN CRITICAL THINKING

Culture plays an important role in shaping the way individuals understand, interpret, and evaluate information. Culture as a symbolically inherited system of meaning, which serves as a framework for individuals to interpret experience. In the context of academic writing, culture influences rhetorical preferences, argumentation structures, and the way ideas are presented [41, 42].

Recent research shows that critical thinking practices are highly contextual and influenced by different cultural norms [43, 44]. For example, in some cultural contexts, explicit arguments and direct criticism are considered inappropriate, thus influencing the way students express their critical thinking. This shows that critical thinking is inseparable from socio-cultural contexts, and that cultural variation can result in variations in academic thinking patterns and writing.

4. THE INTERACTION OF LANGUAGE, CULTURE, AND COGNITION: AN INTEGRATIVE MODEL

Based on the theoretical synthesis presented above, this study proposes a multidimensional construction model of critical thinking in academic writing. This model conceptualizes critical thinking as a result of the dynamic interaction between cognitive reasoning, linguistic mediation, and cultural interpretation [45-48]. In addition, research in the field of second language writing shows that writing performance is the result of an integration between language competence, content knowledge, and cognitive strategies [49, 50]. For this reason, critical thinking cannot be understood as a stand-alone ability, but rather as a product of complex

interactions between dimensions. These dimensions are not independent but interrelated, together shaping how individuals construct, express, and evaluate academic arguments. The structural relationships between these dimensions are illustrated in Figure 1.

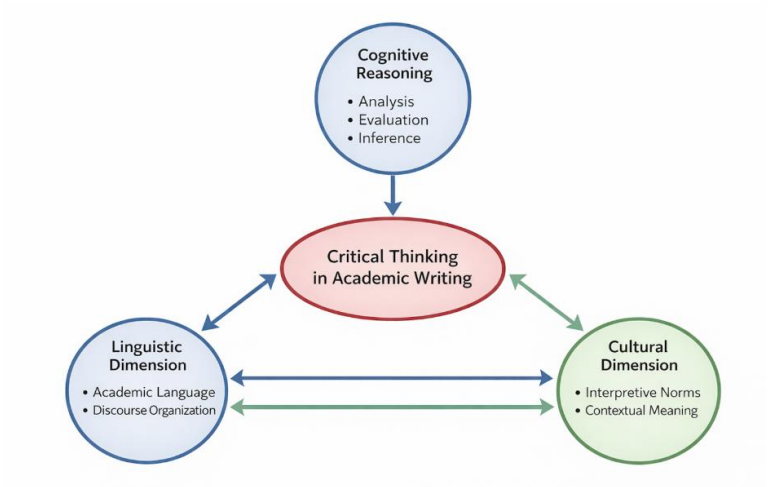

FIGURE 1. Multidimensional construct of critical thinking in academic writing.

As shown in Figure 1, cognitive reasoning provides an analytical foundation for critical thinking, while linguistic mediation allows for the articulation and organization of ideas in academic discourse. At the same time, cultural interpretations shape how meaning is constructed and evaluated in a particular socio-cultural context. The interaction between these dimensions is reciprocal rather than linear, suggesting that critical thinking in academic writing arises from a dynamic and context-dependent process. This multidimensional structure also provides a conceptual basis for applying Rasch-based measurements to examine latent and dimensional interactions in data.

5. MULTIDIMENSIONAL CONSTRUCT MEASUREMENT VIA RASCH MODEL

In the context of measurement, the Rasch model offers a robust probabilistic framework for evaluating the quality of instruments as well as modeling the relationship between individual abilities and the difficulty level of the item [51-53]. One of the main advantages of the Rasch model is its ability to produce invariant measurements, thus allowing for fair comparisons between individuals and items. Nevertheless, the Rasch model classically assumes unidimensionality, i.e. that the instrument measures one major latent construct. In practice, many constructs in education, including critical thinking, are multidimensional, so that there are often indications of misfit, residual dimensionality, and person-item mismatches [52, 54]. Therefore, Rasch analysis serves not only as a tool for evaluating instruments, but also as a means of exploring construct structures.

In this study, the use of the Rasch model allowed the identification: (i) measurement quality (reliability and separation), (ii) item function (fit and misfit), (iii) dimensional structure (residual PCA), (iv) and person-item suitability (Wright map). Thus, this approach makes a methodological contribution to understanding the complexity of critical thinking constructs in academic writing, while strengthening the argument that they are multidimensional and contextual.

III. METHODS

1. MATERIAL

This study uses a quantitative design with a survey approach that is analyzed using the Rasch model to evaluate the quality of measurement as well as the structure of critical thinking constructs in academic

writing. This approach was chosen because the Rasch model allows for invariant parameter estimation and provides more accurate probabilistic-based measurements than the classical [55]. The research was carried out at the Indonesian Language and Literature Education Study Program, Faculty of Teacher Training and Education, University of Muhammadiyah Purwokerto, Indonesia, which has the characteristics of learning based on academic literacy and scientific writing.

The research participants totaled 257 students who were selected using the purposive sampling technique with the following criteria: (1) have taken academic writing courses, (2) have experience in compiling scientific papers, and (3) are in the middle to final semester. The selection of this group was based on the assumption that critical thinking skills develop through academic experience and ongoing writing practice. This sample count has met the recommendations in Rasch's analysis to produce stable parameter estimates, especially in construct measurements with relatively large numbers of items. In addition, the homogeneity of respondents' academic backgrounds allows for a more focused analysis of the relationship between language, culture, and critical thinking.

2. RESEARCH INSTRUMENT

The research instrument consisted of a 36-item Likert-scale questionnaire designed to measure critical thinking in academic writing as a multidimensional construct encompassing cognitive, linguistic, and cultural dimensions. The development of the instrument followed a systematic multi-stage procedure to ensure content validity, construct alignment, and measurement clarity [56-58].

In the first stage, a theoretical framework was established based on key dimensions of critical thinking, including analysis, evaluation, inference, and argument construction, as well as linguistic mediation and cultural interpretation. Based on this framework, an initial pool of items was generated to represent each construct domain. Each item was designed to reflect the interaction between cognitive reasoning, language use, and contextual interpretation in academic writing. In the second stage, content validation was conducted through expert judgment. Three experts in language education and educational measurement were invited to evaluate the items in terms of construct relevance, clarity of wording, and representativeness of the intended dimensions. The experts were provided with a validation rubric and asked to assess each item using predefined criteria. Based on their feedback, several items were revised to improve conceptual alignment and linguistic precision, while ambiguous or overlapping items were refined. In the third stage, a pilot test was conducted with a small group of students to examine item readability, response clarity, and potential misinterpretation. Feedback from the pilot test was used to further refine item wording and ensure that the items were comprehensible and contextually appropriate. In the final stage, the validated instrument was administered to the full sample. The Likert-scale responses were coded from 1 (strongly disagree) to 5 (strongly agree), and the data were analyzed using the Rasch model to evaluate item functioning, reliability, and dimensionality. This systematic development process ensured that the instrument was theoretically grounded, empirically testable, and aligned with the multidimensional construct proposed in this study.

To ensure construct transparency and alignment, the distribution of all items across dimensions and indicators is systematically presented in Table 1. This mapping demonstrates how each questionnaire item corresponds to specific components of the multidimensional construct, including cognitive, linguistic, and cultural dimensions. The table provides a structural overview of how the instrument operationalizes the theoretical model proposed in this study.

Table 1. Construct–item mapping of the instrument.

Dimension	Indicator	Item Codes
Cognitive	Analysis	I01, I02, I03, I04
Cognitive	Evaluation	I05, I06, I07, I08
Cognitive	Inference	I09, I10, I11, I12
Linguistic	Academic Language	I13, I14, I15, I16

Dimension	Indicator	Item Codes
Linguistic	Discourse Organization	I17, I18, I19, I20
Cultural	Interpretation	I21, I22, I23, I24
Cultural	Contextual Meaning	I25, I26, I27, I28
Cultural	Sociocultural Awareness	I29, I30, I31, I32
Cognitive–Linguistic Integration	Argument Construction	I33, I34, I35, I36

3. PROCEDURE

The research procedure is carried out through several systematic stages which include instrument development, data collection, and Rasch model-based analysis. In the early stages, the instruments are compiled based on critical thinking indicators that include analysis, evaluation, inference, and argumentation construction as well as considering linguistic aspects and cultural contexts in academic writing. The instrument then goes through a content by experts in the field of language education and learning evaluation to ensure the suitability of the construct and clarity of the items, as recommended in the development of educational instruments [59].

Furthermore, limited trials were conducted to identify potential ambiguities of items and ensure the readability of the instrument. This stage is important to improve the quality of the data as well as minimize response errors due to item ambiguity. After the instrument is declared feasible, data collection is carried out online through a digital survey platform in the context of learning academic writing. All respondents were given an explanation of the research objectives, data confidentiality guarantees, and the right to voluntary participation before filling in the instrument, in accordance with the principles of educational research ethics [60, 61]. Charging is carried out independently in a controlled time to reduce the potential for external bias.

The collected data is then analyzed using the Rasch model with the help of Winsteps software. The analysis is carried out in stages including: (1) evaluation of reliability and separation index to measure the consistency of the instrument, (2) item analysis to identify items that do not match the model, (3) respondent fit analysis to detect inconsistent response patterns, (4) dimensionality analysis through Principal Component Analysis (PCA) on residual to test the assumption of unidimensionality, and (5) analysis of Wright's maps to evaluate the suitability between respondents' ability distributions and item difficulty [62].

Through this procedure, the research not only evaluates the quality of the instrument, but also explores the latent structure of critical thinking constructs in academic writing comprehensively. This approach allows the identification of relationships between cognitive, linguistic, and cultural dimensions, as well as provides an empirical basis for the development of the conceptual models proposed in this study.

IV. RESULTS

1. MEASUREMENT QUALITY AND INSTRUMENT RELIABILITY

The results of Rasch's analysis showed that the instrument had a moderate measurement quality. The reliability of the person of 0.67 with a separation index of 1.43 indicates that the instrument is able to distinguish about two groups of respondents' abilities. Meanwhile, the item reliability of 0.65 with a separation of 1.35 indicates that the variation in the difficulty level of the item is relatively limited. As shown in Table 2, the model-based reliability value increased significantly, reaching 0.80 for persons and 0.91 for items. These findings suggest that although the variation in empirical scores is relatively limited, latent constructions can still be stably estimated within Rasch's probabilistic framework.

Beyond the summary reliability indices presented in Table 2, the distribution of person and item measures provides further insight into the measurement quality of the instrument. The distribution of person measures indicates that most respondents are concentrated within the moderate ability range, with relatively few respondents at the higher end of the scale. This clustering suggests limited variability in high-level ability, which reduces the instrument's sensitivity in distinguishing more advanced critical thinking performance.

In contrast, the distribution of item difficulty shows that most items are located within a relatively narrow range of low to moderate difficulty levels. This limited spread contributes to the moderate separation indices observed. The imbalance between person ability distribution and item difficulty further indicates a targeting mismatch, which is consistent with the pattern observed in the Wright map analysis.

Table 2. Summary of Rasch measurement statistics.

Indicator	Value	Interpretation
Reliability Person	0,67	Moderate
Reliability Item	0,65	Moderate
Separation Person	1,43	±2 Ability strata
Separation Item	1,35	Limited item variety
Reliability Model (Person)	0,80	Good
Reliability Model (Item)	0,91	Excellent

Source: Winsteps Output

The difference between the empirical reliability and the reliability of the model indicates an imbalance between the distribution of the respondent's ability and the difficulty level of the item. The instrument showed limited sensitivity, especially in highly capable respondents, which indicated a potential ceiling effect. Furthermore, this limited discriminating power leads to an indication of a ceiling effect, where the instrument is less able to distinguish respondents at the upper level of ability. Thus, although the instrument has sufficient measurement stability within the Rasch model, it is necessary to develop items with a more varied level of difficulty to improve measurement accuracy.

2. ITEM FUNCTION AND MISFIT ANALYSIS

Item fit analysis shows that some items experience significant misfits to the Rasch model, indicating a deviation from the probabilistic model's expectations. The most problematic items are I0035, I0036, and I0034, which had Infit Mean Square (MNSQ) values above 5.0. As shown in Table 3, the items exhibit a very high level of mismatch, which indicates an unstable response pattern that is difficult for the model to predict. This extreme misfit value signifies that the item does not fully represent the consistently measured construct.

Table 3. Items with the highest misfit rate.

Item	Logit	Infit MNSQ	Outfit MNSQ	Interpretation
I0035	0,57	7,58	6,93	Extreme misfit
I0036	0,65	6,71	6,45	Extreme misfit
I0034	0,42	5,83	5,30	High misfit
I0033	0,17	5,19	4,76	High misfit
I0032	0,20	4,51	3,27	Moderate Misfit

Source: Output Winsteps.

Methodologically, this condition can be caused by several factors, including: (1) ambiguity in the representation of constructs, (2) unclear or multi-interpreted item formulations, and (3) the existence of more than one latent dimension in one question item. On the other hand, some items are within an acceptable fit range (0.50–1.50), which indicates that the instrument's core components still have adequate internal consistency. This indicates that the instrument is not completely structurally failing, but is experiencing imbalances in certain parts.

These findings suggest that the instrument is not entirely unidimensional. Instead, instruments represent a hybrid construct, in which cognitive, linguistic, and cultural dimensions interact with each other in shaping critical thinking abilities in academic writing. Furthermore, the presence of items with high misfits indicates that some indicators not only measure critical thinking skills, but also the possibility of carrying additional cognitive burdens, such as linguistic complexity or interpretation based on cultural context. Thus, the misfit in this analysis not solely reflects the weaknesses of the instrument, but also opens up the possibility that the measured construct has a more complex multidimensional structure than previously assumed.

3. CONSTRUCT VALIDITY: DIMENSIONALITY ANALYSIS

Dimensionality tests using Principal Component Analysis (PCA) on residual showed that the variance described by the model was only 5.1%, which is well below the limit commonly used to indicate strong unidimensionality. As shown in Table 4, the eigenvalue for the first contrast is 1.98 with a variance contribution of 5.2%, which indicates the presence of additional latent dimensions outside the main construct. In addition, the second and third contrasts also show relatively high values (1.61 and 1.54, respectively), which further strengthens the indication of the presence of multidimensional structures in the data.

Table 4. Residual PCA analysis results.

Components	Eigenvalue	Percentage	Interpretation
Variance explained model	1,93	5,1%	Low
Contrast 1	1,98	5,2%	Secondary dimensions
Contrast 2	1,61	4,2%	Additional structures
Contrast 3	1,54	4,1%	Multidimensional indications

Source: PCA Output

These findings suggest that the assumption of unidimensionality in the measurement of critical thinking ability in academic writing is not fully fulfilled. Empirically, this instrument does not represent just one single construct, but contains several latent dimensions that intersect.

Conceptually, these results indicate that critical thinking abilities in the context of academic writing are multidimensional, which not only reflects cognitive capacity alone, but also involves complex interactions between: (i) analytical reasoning, (ii) linguistic competence in organizing ideas, and (iii) interpretive abilities influenced by social and cultural contexts. Furthermore, the low variance described by the model shows that the construct structure is not completely linear, but rather complex and layered construct. This implies that measurement approaches that assume a single dimension have the potential to simplify more complex cognitive realities in academic writing practice.

Thus, the results of this analysis not only reveal the limitations of the instrument, but also provide a theoretical contribution that critical thinking in academic writing should be understood as an integrative construct based on language and culture, rather than just an isolated cognitive ability. These findings of multidimensionality suggest that critical thinking constructs in academic writing not only represent cognitive capacity, but also contain linguistic and cultural dimensions as latent components that interact with each other.

4. PERSON–ITEM FIT (WRIGHT MAP ANALYSIS)

Analysis of the Wright map in Figure 1 shows the relationship between the distribution of respondent abilities and the difficulty level of items in logit units. In general, the distribution of respondents is concentrated in the low to medium ability range, with the highest density being around −1 to 0 logit values. This indicates that the majority of respondents are at a medium level of ability, with relatively limited variation in ability at high levels.

https://doi.org/10.48161/qaj.v6n3a2444

On the other hand, the distribution of items shows a relatively widespread pattern, but it is dominated by items with low to medium difficulty. Most items are in the range of −1 to +1 logit (easy and moderate categories), which suggests that instruments tend to be more sensitive in measuring capabilities at the basic to intermediate level. However, there are several items that are at a high difficulty level, especially I0035 and I0036, which lie far above the ability distribution of the majority of respondents. The categorization of the difficulty level of an item (easy, moderate, difficult) is based on the distribution of logit relative to the mean of the item (0 logit), which is a common practice in the interpretation of Wright's map in the analysis of Rasch [54, 63].

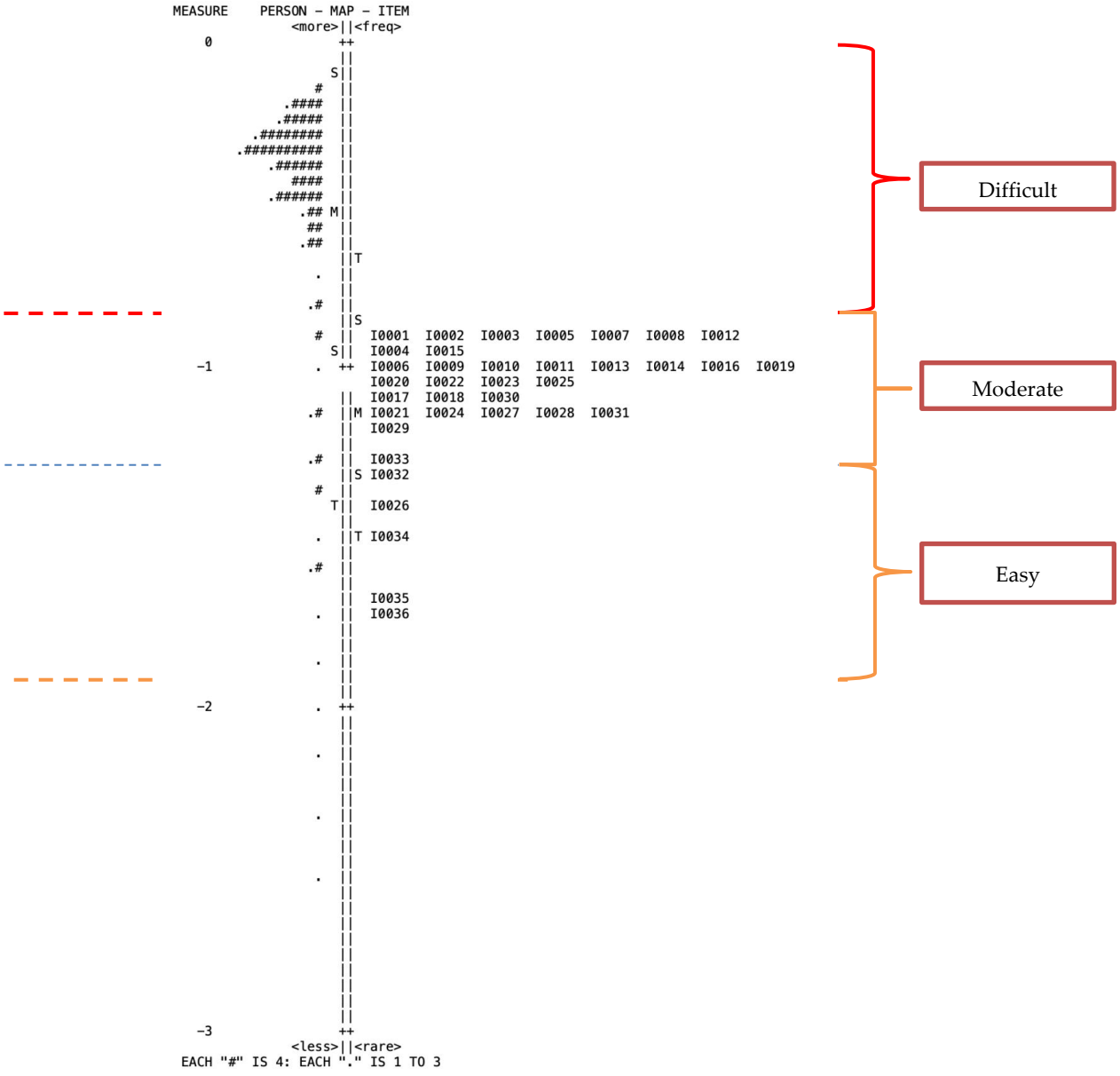

FIGURE 1. Wright Map showing the distribution of person ability and item difficulty on the same logit scale. The left side represents respondents, while the right side represents item difficulty.

The misalignment between the distribution of the respondent's ability and the difficulty level of the item indicates a targeting mismatch. The instrument has not been able to reach the variation of respondents' abilities optimally, especially at high skill levels. This condition is reinforced by the presence of items with extreme difficulty levels that are not offset by the number of respondents with equivalent abilities, so that the measurement information in that range becomes less stable. In addition, the concentration of items in the moderate and easy categories showed that the instrument was more effective in distinguishing respondents at the intermediate ability level, but was less able to provide adequate differentiation in the high-ability group. This has implications for the potential for a ceiling effect, where respondents with high abilities cannot be optimally measured due to the limitations of challenging item variety.

The Wright map visualization provides a clearer representation of the targeting between respondent ability and item difficulty. As shown in Figure 1, the distribution of items is concentrated within the low to moderate logit range, while fewer items represent higher levels of difficulty. This imbalance suggests that the instrument is more sensitive to measuring moderate levels of ability, with limited discrimination at the upper range. The visualization thus supports the presence of a targeting mismatch and reinforces the need for additional high-difficulty items. Overall, the Wright map indicates that although the instrument has fairly good measurement capabilities at the intermediate capability range, further development is needed, particularly through the addition of items with high difficulty levels and extreme recalibration of items. Thus, the compatibility between the distribution of persons and items can be improved, so that the instrument is more sensitive and accurate in measuring critical thinking skills in academic writing comprehensively.

5. RESPONDENT FIT (PERSON FIT)

The analysis showed that not all respondents provided a consistent response pattern to the Rasch model. Based on Infit Mean Square (Infit MNSQ) value, several respondents were found with very high scores (>8.0), which indicated a response that was not predicted by the model. The respondents with the highest misfit rates are presented in Table 5.

Table 5. Respondents with the highest misfit.

Respondent ID	Logit	Infit MNSQ	Outfit MNSQ	Interpretation
0063	-0,27	1,77	9,90	Extreme inconsistencies
0161	-0,27	1,76	9,58	Extreme inconsistencies
0038	-0,71	6,71	9,47	Heavy misfit
0194	-0,37	2,23	8,89	High misfit

Source: Winsteps Output (Person Fit)

Quantitatively, the very high value of the MNSQ Outfit showed that the responses given by the respondents were unpredictable (random noise) and did not match the probabilistic pattern of the Rasch model. This indicates a mismatch between the respondents' latent abilities and the response patterns, which has the potential to reduce the stability of estimates at the individual level.

Interestingly, all respondents who experienced misfit were in the intermediate ability range (−0.71 to −0.27 logits), not at the extreme level. This pattern shows that response inconsistencies are not solely caused by low ability, but are related to the variability of cognitive strategies in responding to items. Respondents at the middle level tend to use inconsistent approaches, resulting in less stable response patterns. This phenomenon can be influenced by several factors, such as the ambiguity of item constructs, response fatigue, and differences in cognitive strategies in understanding and answering items, especially in the context of academic writing that demands a high-level thinking process. These findings indicate that critical thinking skills in academic writing are complex and not homogeneous, but are influenced by the interaction between cognitive, linguistic, and socio-cultural contexts. Therefore, the person fit analysis not only serves as an

indicator of data quality, but also as a diagnostic tool to identify variations in respondents' thinking patterns in the response process.

V. DISCUSSION

1. MEASUREMENT AND INSTRUMENT FINDINGS

This study provides a detailed examination of the measurement quality of an instrument designed to assess critical thinking in academic writing using the Rasch model. The results indicate that the instrument demonstrates acceptable measurement stability within a probabilistic framework, although empirical reliability remains at a moderate level. The person and item reliability indices, together with the relatively low separation values, suggest that the instrument is capable of distinguishing general differences in respondent ability but is less effective in differentiating higher levels of performance.

The identification of misfitting items further indicates that not all items align consistently with the underlying construct. This misalignment may reflect ambiguity in item formulation or the presence of overlapping dimensions embedded within individual items. In addition, the mismatch between the distribution of respondent ability and item difficulty, as observed in the Wright map, points to a targeting issue in which the instrument is more sensitive to moderate levels of ability but less capable of capturing advanced critical thinking performance. Such a pattern has also been reported in Rasch-based studies in educational contexts, where limited item spread reduces measurement precision at the upper ability range [64]. Taken together, these findings suggest that while the instrument is functionally adequate, further refinement is required. In particular, expanding the range of item difficulty and improving item clarity would enhance the instrument's discriminative capacity and alignment with the intended construct.

2. THEORETICAL IMPLICATIONS

Beyond measurement considerations, the findings offer important implications for the conceptualization of critical thinking in academic writing. The presence of residual dimensions and item misfit patterns indicates that critical thinking cannot be adequately represented as a unidimensional cognitive construct. Instead, the results support a multidimensional perspective in which cognitive reasoning interacts with linguistic mediation and cultural interpretation.

This interpretation is consistent with sociocultural perspectives that view thinking as mediated by language and shaped by contextual factors [65-67]. In academic writing, critical thinking is not only reflected in analytical reasoning but also in the ability to organize ideas through language and interpret meaning within specific cultural frameworks. The observed variation in response patterns, particularly among respondents in the intermediate ability range, further suggests that critical thinking performance is influenced by differences in linguistic competence and interpretive strategies. Accordingly, the findings reinforce the argument that critical thinking in academic writing should be understood as an integrative construct. Rather than functioning as an isolated cognitive skill, it emerges from the interaction between reasoning processes, language use, and sociocultural context. This perspective extends existing models of critical thinking by situating it within academic discourse practices and providing empirical support for its multidimensional nature.

3. INSTRUMENT LIMITATIONS AND FUTURE DIRECTIONS

Despite its contributions, this study has several limitations related to the measurement instrument. First, the relatively narrow range of item difficulty restricts the instrument's ability to discriminate among respondents with higher levels of critical thinking ability. Second, the presence of misfitting items suggests that some items may require revision to improve their alignment with the intended construct. Third, the study was conducted within a specific educational context, which may limit the generalizability of the findings to other populations.

These limitations point to several directions for future research. Further studies should focus on refining the instrument by developing items with greater variation in difficulty and clearer construct representation. In addition, the application of more advanced measurement approaches, such as multidimensional Rasch

models, would allow for a more precise examination of the interaction between cognitive, linguistic, and cultural dimensions. Expanding the study across diverse educational settings and incorporating qualitative approaches, such as discourse analysis, may also provide deeper insight into the processes underlying critical thinking in academic writing.

VI. CONCLUSION

This study provides empirical insights into the measurement and conceptualization of critical thinking in academic writing within a Rasch measurement framework. The findings indicate that the instrument demonstrates moderate empirical reliability alongside stable model-based reliability, suggesting consistent estimation within a probabilistic framework. Item fit analysis identified several misfitting items, indicating potential additional cognitive or linguistic demands, while residual PCA revealed indications of additional latent structures. Furthermore, the Wright map analysis showed a mismatch between respondent ability distribution and item difficulty levels. Collectively, these results suggest that critical thinking in academic writing may not be fully captured as a unidimensional construct and may involve interactions between analytical reasoning, linguistic competence, and sociocultural interpretation.

Methodologically, this study highlights the value of Rasch-based analysis for evaluating instrument quality and exploring latent construct structures in educational research. Pedagogically, the findings underscore the importance of integrating cognitive, linguistic, and cultural dimensions in academic writing instruction. However, these findings should be interpreted as preliminary, and further confirmatory analysis is required to establish the multidimensional structure of the construct. The study is also limited by the relatively homogeneous sample and single institutional context. Future research should involve more diverse populations, apply advanced measurement models such as multidimensional Rasch or confirmatory approaches, and incorporate mixed-method designs to further examine the structure of critical thinking in academic writing.

Funding Statement

The authors received no financial support for this research.

Author Contributions

Conceptualization, K.S. and M.T.A.; methodology, O.Y.P. and M.T.A.; software, O.Y.P. and H.; validation, K.S.; H. and O.Y.P.; formal analysis, K.S.; investigation, K.S. and O.Y.P.; resources, A.D.; data curation, O.Y.P.; writing—original draft preparation, K.S. and O.Y.P.; writing—review and editing, M.T.A.; project administration, K.S. and A.M. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

The authors declare that there is no conflict of interest.

Data Availability Statement

The data presented in this study are available on request from the corresponding author.

Acknowledgments

Not applicable.

REFERENCES

1. Hao, L., Yusoff, S. M., & Tian, K. (**2026**). Impacts of AI-enhanced task-based learning on EFL postgraduates' higher order thinking skills and English academic writing self-efficacy. *Journal of English for Academic Purposes, 80*, 101652.
2. Yin, X., Saad, M. R. B. M., & Halim, H. B. A. (**2023**). A systematic review of critical thinking instructional pedagogies in EFL writing: What do we know from a decade of research. *Thinking Skills and Creativity, 49*, 101363.
3. Manalo, E., & Sheppard, C. (**2016**). How might language affect critical thinking performance? *Thinking Skills and Creativity, 21*, 41–

49.
4. Mores, H., & Pradipto, Y. D. (**2025**). The Role of Cultural Norms and Religious Values in a Moderated-Mediation Model of Inconspicuous Consumption Purchase Intention of Luxury Lingerie. *Qubahan Academic Journal*, *5*(4), 292-321.
5. Ryan, U. (**2022**). Solidarity acts in the context of language diversity: Students' encounters with epistemological difference. *Journal of Mathematical Behavior, 66*, 100946.
6. Kuleshova, S., & Pleyer, M. (**2025**). From extant to extinct: The role of experiments and interdisciplinary inferences in studying cognitive and language evolution. *Journal of Archaeological Science, 184*, 106423.
7. Siem, T. T., Nguyen, H. T. T., My, D. H., To, A. T., Phong, T. K., & Van, T. T. T. (**2026**). From branding to belonging: How employer image and social identity shape organizational attractiveness. *Social Sciences & Humanities Open, 13*, 102429.
8. Kriener, F., Streeb, M. N., Fischer, S., & Baum, M. (**2025**). Anticipating the future: Job seeker's evaluation of employer attractiveness during organizational change. *European Management Journal*.
9. Liu, L., & Shen, H. (**2026**). Fostering critical thinking through STEM undergraduate research: Mechanisms and challenges in a Chinese research university. *Thinking Skills and Creativity, 61*, 102173.
10. Park, J. H., Myszkowski, N., & Niu, W. (**2026**). Exploring the intersection of creativity and critical thinking among students: A meta-analysis. *Thinking Skills and Creativity, 60*, 102119.
11. Haryadi, R., Pujiastuti, H., & Utomo, D. W. (**2026**). Interactive visualization of thermodynamic concepts through augmented reality to improve critical thinking. *Computers & Education: X Reality, 8*, 100137.
12. Loginov, A. (**2025**). Digital metacognition: Awareness, regulation, and analysis of thinking in digital environments. *Next Research, 2*(4), 100770.
13. Applin, S. (**2024**). Language in critical thinking discourses. In *Reference Module in Social Sciences*. Elsevier.
14. Tung, C.-A. J. (**2026**). Integrating visible thinking routines into an EFL writing class. *Thinking Skills and Creativity, 61*, 102141.
15. Zubizarreta, D., Beccia, A. L., & Platt, J. M. (**2026**). Quantitative intersectionality research: Past, present, and future directions. *Annals of Epidemiology*, 110061.
16. Brown, D. M. Y., Meca, A., & Osman, A. (**2026**). Development of the multidimensional inventory of physical activity identity. *Psychology of Sport and Exercise, 82*, 102990.
17. Lavenne-Collot, N., Planche, P., & Vaivre-Douret, L. (**2025**). Empathy in subjects with high intellectual potential. *Intelligence, 112*, 101935.
18. Tamano, H., Hino, H., & Mochihashi, D. (**2025**). Misspecifying non-compensatory as compensatory IRT. *Behaviormetrika, 52*(2), 393–412.
19. Reckase, M. D. (**2018**). Logistic multidimensional models. In *Handbook of Item Response Theory*. CRC Press.
20. Hartig, J., & Höhler, J. (**2009**). Multidimensional IRT models for competencies. *Studies in Educational Evaluation, 35*(2–3), 57–63.
21. Bravo, L., Gutierrez, C. R., Cobo, M., & Figueredo, M. (**2025**). Multiscale modelling in ammonia synthesis. *Case Studies in Chemical and Environmental Engineering, 12*, 101299.
22. Li, M., & Wang, F. (**2026**). EAL scholars' experiences of writing for publication in English. *English for Specific Purposes, 81*, 16–31.
23. Awolesi, O., Ghafari, F., & Reams, M. (**2025**). Indoor environmental quality assessment. *Energy and Built Environment*.
24. Celik, S. U. (**2025**). Integrating artificial intelligence into scientific writing. *American Journal of Surgery, 250*, 116657.
25. Dominguez Romero, E., & Bobkina, J. (**2026**). Visual critical thinking in EFL instruction. *Thinking Skills and Creativity, 60*, 102054.
26. Sun, Y., Wang, H., Fan, Y., Wang, F., & Tang, X. (**2026**). ICT and creative thinking. *Thinking Skills and Creativity, 59*, 102004.
27. Astaíza Martínez, A. F., & Rojas León, G. A. (**2024**). Systems thinking and citizenship education. *Kybernetes, 54*(14), 7573–7590.
28. Şendağ, S., Caner, M., Gedik, N., & Toker, S. (**2024**). Mobile listening and metacognition. *Thinking Skills and Creativity, 54*, 101656.
29. Matsuoka, M., Manalo, E., Rosário, P., & Núñez, J. C. (**2026**). Narratives for learning development. *Thinking Skills and Creativity, 61*, 102114.
30. Sun, Q., Zhang, L. J., & Carter, S. (**2024**). Metacognitive experiences in EFL writing. *Learning and Individual Differences, 115*, 102540.
31. Lantolf, J. P. (**2024**). Sociocultural theory and language education. In *Reference Module in Social Sciences*. Elsevier.
32. Huang, K. (**2026**). Teacher support and learning engagement. *Acta Psychologica, 263*, 106380.
33. Lobaccaro, L. (**2022**). Cognitive semiotics. *Zeitschrift für Semiotik, 44*(3–4), 157–182.
34. Kravchenko, A. V. (**2024**). Language awareness. *Language Sciences, 105*, 101659.
35. Gahrn-Andersen, R. (**2025**). Enlanguaged affordances. *Language & Communication, 104*, 121–130.
36. Takatsu, H., Suzuki, S., Eguchi, M., Matsuura, R., Saeki, M., & Matsuyama, Y. (**2026**). Oral proficiency assessment model. *Computer Speech & Language, 95*, 101860.

37. Wu, H., Liu, W., & Zeng, Y. (**2024**). AI-assisted L2 learning attitude scale. *Acta Psychologica, 248*, 104376.
38. Kalinowski, E., Egert, F., Gronostaj, A., & Vock, M. (**2020**). Academic language proficiency development. *Teaching and Teacher Education, 88*, 102971.
39. Luo, N., Zhang, Q., Hong, C., Zhu, J., & Wang, G. (**2026**). Linguistic decision-making models. *Information Sciences, 730*, 122905.
40. Liu, Z., Yi, H., Yao, Y., & Wang, J. (**2025**). Three-way decision model. *Applied Soft Computing, 184*, 113816.
41. Lei, H. (**2026**). Rhetorical shifts in AI age. *Linguistics and Education, 92*, 101508.
42. Meng, K., Koh, C., Zheng, Z., Ba, Z., & Song, M. (**2025**). Citation behavior and rhetoric. *Journal of Informetrics, 19*(4), 101729.
43. Rezaee, A. A., & Saleh, E. (**2025**). Critical thinking in EFL textbooks. *Thinking Skills and Creativity, 57*, 101782.
44. Verma, P., Singh, A., & Kumar, V. (**2026**). Strategic thinking and culture. *Strategic Business Research, 2*(1), 100048.
45. Zheng, Y., Yu, S., Liu, C., & Jiang, L. (**2022**). L2 writing teacher cognition. *System, 109*, 102870.
46. Truong, K. D., Cong-Lem, N., & Li, B. (**2025**). Language teacher identity. *Teaching and Teacher Education, 158*, 104967.
47. Cui, H., Mahfoodh, O. H. A., Dong, H., & Ulanbek, M. (**2026**). Human-AI competence in writing. *System, 137*, 103937.
48. Du, Y., & Reynolds, B. L. (**2025**). Chatbot affordances in language learning. *Acta Psychologica, 259*, 105307.
49. Zhu, D. (**2026**). Neuro-symbolic cognition framework. *Applied Soft Computing, 193*, 114887.
50. Cacciante, L., & Ranzini, M. (**2026**). Language and cognition integration. *Brain and Language, 273*, 105701.
51. Suherman, S., & Vidákovich, T. (**2025**). Creativity inventory Rasch analysis. *Journal of Creativity, 35*(3), 100110.
52. DeMars, C. E. (**2025**). Rasch models. In *Reference Module in Social Sciences*. Elsevier.
53. Sarifah, I., et al. (**2025**). Rasch model in educational policy. *Social Sciences & Humanities Open, 12*, 101752.
54. Verdú-Soriano, J., & González-de la Torre, H. (**2024**). Rasch analysis in nursing. *Enfermería Clínica*, 34*(6), 493–506.
55. Chakraburty, S., et al. (**2025**). Rasch analysis of AI learning concepts. *Information Learning Sciences, 126*(78), 445–471.
56. Abdeta, M. T., Olamo, T. G., & Woldeyes, K. M. (**2026**). Collaborative writing instruction. *Acta Psychologica, 263*, 106307.
57. Alkamel, M. A. A., & Alwagieh, N. A. S. (**2024**). ChatGPT in academic writing. *Social Sciences & Humanities Open, 10*, 101095.
58. Li, X. (**2026**). Cognitive mechanisms in translation. *Acta Psychologica, 263*, 106215.
59. Han, C. (**2026**). AI evaluation of translation. *Research Methods in Applied Linguistics, 5*(1), 100300.
60. Austin, T., & Medina Riveros, R. A. (**2025**). Ethics in language research. *Research Methods in Applied Linguistics, 4*(2), 100221.
61. Wiese, L. J., Patil, I., Schiff, D. S., & Magana, A. J. (**2025**). AI ethics education review. *Computers and Education: Artificial Intelligence, 8*, 100405.
62. Christensen, K. S., & Ammentorp, J. (**2024**). SE-12 Rasch analysis. *PEC Innovation, 4*, 100296.
63. Davis, D. R., & Boone, W. (**2021**). OLIW scale Rasch analysis. *International Journal of Educational Research Open, 2*, 100054.
64. Riemenschneider, M., Trénel, P., Nørgaard, M., & Boesen, F. (**2022**). Fatigue impact scale validation. *Multiple Sclerosis and Related Disorders, 65*, 104012.
65. Zhang, A., & Xie, J. (**2025**). Linguistic landscape tourism. *Annals of Tourism Research, 115*, 104019.
66. Aldayel, H. S., Pravitasari, H., & Fadillah, N. (**2026**). EFL writing difficulties. *Acta Psychologica, 263*, 106250.
67. Morgan-Thomas, A. (**2025**). Genre of power in academia. *European Management Journal*.