



Developing and Validating a 4D Reflection Rubric for Assessing Teacher Professional Reflection in Narrative Writing

Lilis Ernawati*, **Ribut W. Eriyanti**, **Masduki**, **Siti Agustin**, and **Apriani Riyanti**

To cite this article: L. Ernawati, R. W. Eriyanti, S. Agustin, and A. Riyanti, “Developing and Validating a 4D Reflection Rubric for Assessing Teacher Professional Reflection in Narrative Writing,” *Women, Educ. Soc. Welf.*, vol. 3, no. 1, pp. 177–196, 2026. <https://doi.org/10.70211/wesw.v3i1.440>



Published online: March 30, 2026



Submit your article to this journal



View crossmark data



Developing and Validating a 4D Reflection Rubric for Assessing Teacher Professional Reflection in Narrative Writing

Lilis Ernawati*, Ribut W. Eriyanti, Masduki, Siti Agustin and Apriani Riyanti

Received: December 13, 2025

Revised: January 16, 2026

Accepted: March 25, 2026

Online: March 30, 2026

Abstract

Assessing teacher reflection in narrative texts remains challenging because existing rubrics often privilege explicit analytical writing and may overlook reflective meanings embedded in stories. This study aimed to develop and validate a multidimensional, genre-sensitive 4D Reflection Rubric for assessing pedagogical reflection in teacher-authored short stories. Using a Research and Development design adapted from the 4D model, the study involved 30 in-service elementary school teachers who produced classroom-based narratives, five expert validators, and five independent raters. The rubric was developed across four dimensions: cognitive, affective-relational, ethical, and action-oriented reflection. Data were analyzed using the Content Validity Ratio, Aiken's V, Intraclass Correlation Coefficient, and descriptive statistics. The findings showed that nine of twelve indicators achieved full expert agreement, while three indicators required refinement. All indicators demonstrated satisfactory clarity, with Aiken's V values ranging from 0.80 to 0.93. The rubric also showed excellent inter-rater reliability, with an ICC value of 0.90. Ethical reflection emerged as the most prominent dimension, followed by action-oriented and affective-relational reflection, whereas cognitive reflection was less explicitly articulated. These findings indicate that the 4D Reflection Rubric is valid, reliable, and responsive to both explicit and implicit forms of teacher reflection. This study contributes a genre-sensitive assessment tool for strengthening teacher professional development, feedback quality, and institutional capacity building in educational settings.

Keywords: Pedagogical Reflection; Narrative Assessment; Rubric Development; Teacher Professional Development; Content Validity.

Publisher's Note:

WISE Pendidikan Indonesia stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright:

©

2026 by the author(s).

License WISE Pendidikan Indonesia, Bandar Lampung, Indonesia. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY 4.0) license

(<https://creativecommons.org/licenses/by/4.0/>).



INTRODUCTION

Improving the quality of education requires sustained attention to teacher professionalism, institutional capacity, and the quality of pedagogical services provided in schools [1], [2], [3]. In contemporary education systems, teachers are not merely curriculum implementers but reflective professionals who interpret classroom realities, respond to learners' needs, negotiate ethical responsibilities, and make instructional decisions that shape students' learning experiences. Therefore, teacher reflection is increasingly viewed as a critical mechanism for strengthening professional learning, improving instructional quality, and supporting school-level capacity for continuous improvement [4], [5], [6].

Pedagogical reflection refers to teachers' ability to examine classroom experiences, evaluate instructional decisions, understand emotional and relational dynamics, and formulate future improvements based on professional judgment. Reflection is not limited to cognitive reasoning; it also involves affective sensitivity, ethical awareness, and action-oriented commitment. In this study, pedagogical reflection is conceptualized as a multidimensional construct consisting of cognitive reflection, affective-relational reflection, ethical reflection, and action-oriented reflection. This multidimensional perspective recognizes teaching as a complex professional practice that involves thinking, feeling, valuing, and acting within specific classroom and institutional contexts.

The importance of reflective practice in teacher professional development has been widely discussed in educational literature. Prayogi et al. [7] conceptualized reflection as a disciplined form of inquiry grounded in experience, while Marshall et al. [8] emphasized reflection as a professional process through which practitioners make sense of uncertain and complex situations. Later studies have further shown that reflective practice contributes to instructional improvement, professional growth, and teachers' capacity to respond to diverse classroom challenges [9], [10], [11]. However, previous research also indicates that teachers' reflective writing often remains at a descriptive level, focusing mainly on recounting classroom events rather than critically examining pedagogical assumptions, dilemmas, or alternative actions [12], [13], [14]. This condition raises an important methodological question regarding whether teachers' reflection is genuinely limited or whether existing assessment tools are not sufficiently sensitive to the forms through which teachers express reflection.

This issue becomes especially significant when reflection is expressed through narrative texts. Teachers often make sense of their professional experiences through stories, anecdotes, and classroom narratives. Narrative writing allows teachers to represent pedagogical challenges through events, characters, dialogue, emotional responses, and resolutions. Unlike analytical essays, narrative texts do not always express reflection through explicit argumentation or direct conceptual explanation. Reflective meaning may instead be embedded implicitly in the development of the story, the teacher's response to classroom conflict, or the moral and pedagogical resolution of a teaching experience. As Trzebiński et al. [15] and Owiti et al. [16] suggest, narratives are not merely records of events but meaningful forms through which individuals organize experience and construct professional understanding.

Despite the growing recognition of narrative as a meaningful form of teacher reflection, the assessment of reflective practice remains largely influenced by frameworks designed for analytical or expository writing. Existing reflection rubrics have contributed significantly to

distinguishing levels of reflection, ranging from descriptive accounts to critical or transformative reflection [17], [18], [19]. However, many of these frameworks tend to privilege explicit reasoning, structured argumentation, and direct articulation of insight. This creates a potential mismatch between how teachers express reflection in narrative forms and how their reflection is assessed. From the perspective of assessment validity, such a mismatch is problematic because an assessment instrument should measure the intended construct in a manner that is appropriate to the context, task, and form of performance being evaluated [20], [21], [22].

In the Indonesian educational context, the need for a more context-sensitive reflection assessment tool is increasingly relevant [23], [24], [25]. Current educational reform emphasizes student-centered learning, character formation, contextual pedagogy, and continuous teacher professional learning. These expectations require teachers to reflect not only on instructional techniques but also on learner diversity, classroom relationships, ethical decision-making, and future pedagogical improvement. In this regard, a valid and reliable rubric for assessing teacher reflection can support professional development by helping teachers and institutions identify the depth, quality, and direction of reflective practice. Such an instrument is also relevant to education systems and social welfare services because it contributes to workforce professionalism, service quality, and institutional capacity building in schools.

Responding to this need, the present study develops and validates a 4D Reflection Rubric for assessing teacher-authored short stories. The rubric is designed to capture four dimensions of pedagogical reflection: cognitive, affective-relational, ethical, and action-oriented reflection. Unlike conventional rubrics that mainly assess explicit analytical statements, this rubric adopts a genre-sensitive orientation by recognizing that reflection in narrative texts may be expressed both explicitly and implicitly. Therefore, the rubric is intended to provide a more nuanced and context-responsive framework for evaluating reflective competence in teacher-authored narratives.

The novelty of this study lies in its integration of multidimensional pedagogical reflection with narrative-sensitive assessment. Rather than treating teacher-authored stories as informal or less rigorous forms of reflection, this study positions narrative writing as a professional medium through which teachers interpret classroom realities, negotiate values, and construct pedagogical insight. The study therefore contributes to the development of assessment tools that are not only psychometrically sound but also responsive to the communicative forms through which teachers represent professional experience.

Accordingly, this study aims to develop and empirically validate a multidimensional 4D Reflection Rubric for teacher-authored short stories. The study examines the dimensions that constitute pedagogical reflection in teacher-authored narrative texts, explains how these dimensions can be operationalized into a valid and reliable rubric, and evaluates the extent to which the rubric provides consistent and meaningful differentiation of reflection levels in teacher-authored narratives.

METHODS

This study employed a Research and Development (R&D) design adapted from the 4D model proposed by Thiagarajan et al. [26], consisting of four stages: Define, Design, Develop, and Disseminate. The study adopted a mixed-methods orientation, combining qualitative input from expert judgment with quantitative analysis for validity and reliability testing. This combination was considered necessary, as the development of an assessment rubric requires both conceptual grounding and empirical verification to ensure its usability in real contexts. The overall research procedure is illustrated in Figure 1.

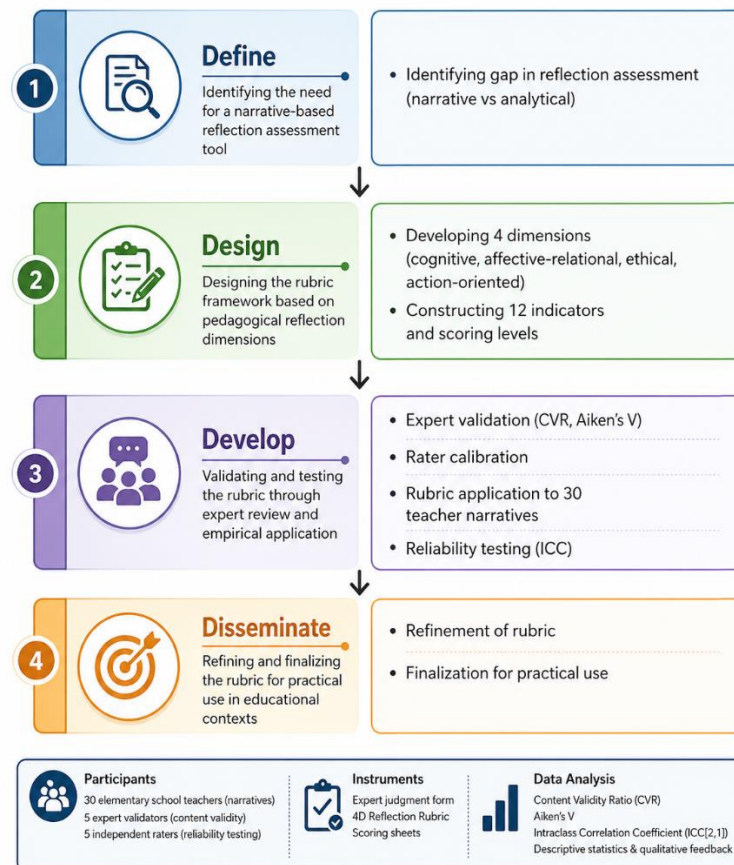


Figure 1. Research Procedure Based on the 4D Model

Research Design

This study employed a Research and Development design adapted from the 4D model proposed by Thiagarajan et al. [26], which consists of four interconnected stages: define, design, develop, and disseminate. This design was selected because the study aimed not only to examine an existing phenomenon but also to produce and validate an assessment instrument that can be used in teacher professional development contexts. The development of the rubric was therefore positioned as both a methodological and practical response to the need for a valid, reliable, and context-sensitive tool for assessing pedagogical reflection in teacher-authored narrative texts.

The define stage focused on identifying the conceptual and practical need for a narrative-based reflection assessment rubric. This stage involved examining the gap between the

expected depth of teachers' pedagogical reflection and the limitations of existing reflection rubrics that tend to privilege explicit analytical writing. The conceptual foundation of the rubric was also established at this stage by defining pedagogical reflection as a multidimensional construct comprising cognitive, affective-relational, ethical, and action-oriented dimensions.

The design stage involved constructing the initial rubric framework. The four dimensions of pedagogical reflection were operationalized into twelve indicators, each accompanied by progressive performance descriptors. The descriptors were arranged across four levels of reflection: descriptive, emerging, analytical, and actionable. These levels were designed to distinguish surface-level narration from deeper forms of reflective engagement while remaining sensitive to narrative texts, in which reflection may be expressed implicitly through classroom events, teacher responses, emotional tensions, ethical considerations, and story resolutions.

The develop stage involved expert validation, revision, rater calibration, and empirical testing. Expert validators evaluated the relevance and clarity of the rubric indicators, while independent raters applied the rubric to teacher-authored short stories. Prior to the scoring process, raters participated in a calibration session to establish a shared understanding of the rubric descriptors, particularly in identifying implicit forms of reflection embedded in narrative structures.

The disseminate stage was limited to the refinement and documentation of the validated rubric prototype. Although broader dissemination was not conducted in this study, the final rubric was prepared as a usable instrument for assessing teacher reflection in narrative writing. This stage strengthened the practical contribution of the study by ensuring that the rubric was not only theoretically grounded but also applicable for professional learning, reflective writing assessment, and institutional capacity building in educational settings.

Instrument

The primary instrument used in this study was the 4D Reflection Rubric, which was developed to assess pedagogical reflection in teacher-authored short stories. The rubric was constructed to capture reflection as a multidimensional construct consisting of cognitive, affective-relational, ethical, and action-oriented reflection. Each dimension was represented by three indicators, resulting in twelve indicators in total. The rubric was designed to assess not only explicit reflective statements but also implicit forms of reflection embedded in classroom events, teacher responses, dialogue, emotional tensions, ethical considerations, and story resolutions.

Table 1. Structure of the 4D Reflection Rubric

Dimension	Indicator Code	Operational Focus	Evidence in Narrative Texts
Cognitive reflection	C1–C3	Identification of pedagogical issues, interpretation of classroom events, and reasoning behind instructional decisions	The teacher identifies a classroom problem, explains why the situation occurred, or justifies a pedagogical response.

Dimension	Indicator Code	Operational Focus	Evidence in Narrative Texts
Affective-relational reflection	AR1–AR3	Recognition of emotions, relational dynamics, empathy, and teacher–student interaction	The narrative shows emotional awareness, sensitivity to students’ needs, or relational negotiation in classroom situations.
Ethical reflection	E1–E3	Awareness of values, fairness, responsibility, care, and professional judgment	The teacher demonstrates concern for justice, inclusion, moral responsibility, or the well-being of learners.
Action-oriented reflection	A1–A3	Formulation of future improvement, pedagogical alternatives, and practical follow-up	The narrative presents lessons learned, planned actions, or strategies for improving future teaching practice.

The rubric applied four progressive performance levels: descriptive, emerging, analytical, and actionable reflection. These levels were developed to distinguish basic narration from deeper reflective engagement. The scoring levels were also intended to support raters in identifying the extent to which teachers moved from merely reporting classroom events toward interpreting, evaluating, and transforming those experiences into future pedagogical action.

Table 2. Scoring Levels of the 4D Reflection Rubric

Score	Reflection Level	General Descriptor	Interpretation in Teacher-Authored Narratives
1	Descriptive reflection	The text mainly recounts classroom events with limited interpretation or evaluation.	The teacher describes what happened but provides little explanation of causes, meanings, values, or future implications.
2	Emerging reflection	The text begins to show awareness of pedagogical meaning, but the reflection remains partial or underdeveloped.	The teacher recognizes a classroom issue or emotional response but does not yet analyze it in depth.
3	Analytical reflection	The text demonstrates interpretation, explanation, and evaluation of classroom experiences.	The teacher examines causes, consequences, relationships, values, or pedagogical considerations within the story.
4	Actionable reflection	The text connects reflection with future pedagogical improvement or professional action.	The teacher formulates clear lessons learned, alternative strategies, or future instructional decisions.

In addition to the 4D Reflection Rubric, this study used two supporting instruments: an expert judgment form and rater scoring sheets. The expert judgment form was used to evaluate the relevance and clarity of each rubric indicator, while the scoring sheets were used by independent raters to record their assessment of teacher-authored short stories. These

supporting instruments ensured that the rubric development process included both expert-based validation and empirical application in an authentic assessment context.

Table 3. Supporting Instruments Used in the Study

Instrument	User	Function	Data Produced
4D Reflection Rubric	Independent raters	To assess the level and dimension of pedagogical reflection in teacher-authored short stories	Reflection scores across four dimensions and four performance levels
Expert judgment form	Expert validators	To evaluate the relevance and clarity of rubric indicators and descriptors	Expert ratings for CVR and Aiken's V analysis
Rater scoring sheet	Independent raters	To record individual rater scores systematically for each narrative text	Score matrix for inter-rater reliability and descriptive analysis
Rater calibration materials	Independent raters	To align raters' understanding of dimensions, indicators, and performance levels	Shared interpretation of scoring criteria before formal assessment

Data Analysis Techniques

Data analysis combined quantitative and qualitative procedures to examine the validity, reliability, and functionality of the 4D Reflection Rubric. Quantitative analysis was used to evaluate content validity, indicator clarity, inter-rater reliability, and score distribution. Qualitative analysis was used to interpret expert comments and refine indicators that were considered conceptually or linguistically ambiguous. This combination was necessary because rubric development requires both psychometric evidence and interpretive refinement to ensure that the instrument is theoretically sound, practically usable, and appropriate for assessing narrative-based reflection.

Table 4. Summary of Data Analysis Procedures

Data Source	Analytical Technique	Purpose of Analysis	Expected Output
Expert ratings on indicator relevance	Content Validity Ratio	To examine whether each indicator was considered essential for assessing pedagogical reflection	Evidence of content relevance for each rubric indicator
Expert ratings on indicator clarity	Aiken's V	To examine the clarity and interpretability of rubric indicators and descriptors	Evidence of clarity and readability of the rubric components
Rater scores on teacher-authored narratives	Intraclass Correlation Coefficient, ICC(2,1)	To examine the consistency of scores across independent raters	Evidence of inter-rater reliability
Reflection scores across dimensions	Descriptive statistics	To summarize total scores, mean scores, and score distribution across reflection dimensions	Profile of reflection patterns in teacher-authored narratives
Expert written comments	Qualitative interpretive analysis	To identify indicators requiring conceptual, linguistic, or operational refinement	Revised rubric indicators and descriptors

Content validity was examined using the Content Validity Ratio proposed by Lawshe (1975). This analysis was used to determine the extent to which expert validators judged each indicator as essential for assessing pedagogical reflection. Aiken's V was then used to assess the clarity of each indicator and descriptor based on expert ratings (Aiken, 1985). The use of both procedures provided complementary evidence: CVR focused on the relevance of the indicators, whereas Aiken's V focused on the clarity and interpretability of the rubric components.

Inter-rater reliability was analyzed using the Intraclass Correlation Coefficient, specifically ICC(2,1). This statistic was selected because the study involved multiple independent raters who assessed the same set of teacher-authored narratives. The ICC analysis was used to determine whether the rubric could produce consistent scores across raters, particularly because narrative reflection involves interpretive judgment and may contain implicit reflective meaning.

Descriptive statistics were used to summarize the distribution of reflection scores across the four dimensions of the rubric. The analysis included total scores, mean scores, and the frequency of reflection levels. These descriptive results were used to identify which dimensions of reflection appeared more prominently in teacher-authored short stories and to determine whether the rubric could meaningfully differentiate levels of reflective engagement.

Table 5. Interpretation Framework for Quantitative Analysis

Analysis Component	Statistical Index	Interpretation Focus	Use in the Study
Content relevance	CVR	Higher values indicate stronger expert agreement regarding the essentiality of an indicator.	To determine whether each rubric indicator adequately represented pedagogical reflection.
Indicator clarity	Aiken's V	Higher values indicate stronger expert agreement regarding the clarity of an indicator or descriptor.	To determine whether rubric descriptors were understandable and usable for assessment.
Scoring consistency	ICC(2,1)	Higher values indicate stronger agreement among independent raters.	To evaluate the reliability of rubric-based scoring across raters.
Reflection profile	Mean, total score, frequency, and percentage	Score patterns indicate the distribution and dominance of reflection dimensions and levels.	To describe how pedagogical reflection appeared across teacher-authored narratives.
Rubric refinement	Expert comments	Repeated concerns indicate indicators requiring revision.	To refine indicators, especially those related to implicit affective-relational and ethical reflection.

Qualitative feedback from expert validators was analyzed interpretively to refine the rubric. Particular attention was given to indicators that received lower agreement or clarity ratings. This refinement process was especially important for the affective-relational and ethical dimensions because these aspects of reflection are often expressed implicitly in

narrative texts. The final interpretation of the rubric’s quality was therefore based not only on statistical results but also on the extent to which the indicators were conceptually coherent, practically assessable, and sensitive to the narrative mode of teacher reflection.

RESULTS AND DISCUSSION

Results

Dimensions of Pedagogical Reflection in Teacher-Authored Narratives

The findings indicate that pedagogical reflection in teacher-authored short stories can be systematically understood through four interrelated dimensions: cognitive reflection, affective-relational reflection, ethical reflection, and action-oriented reflection. These dimensions provide an analytical framework for examining how teachers interpret classroom experiences, negotiate pedagogical challenges, respond to emotional and relational dynamics, and formulate future-oriented professional actions.

The operationalization of these dimensions into twelve indicators enabled the rubric to examine reflection in a more structured and differentiated manner. At the same time, the rubric remained sensitive to the narrative nature of the data, in which reflection was not always expressed through explicit analytical statements but was often embedded in classroom events, teacher–student interactions, emotional responses, moral tensions, and story resolutions. This finding confirms that teacher-authored narratives can serve as meaningful reflective texts when assessed using a genre-sensitive instrument.

Content Validity of the 4D Reflection Rubric

The content validity analysis was conducted to examine the extent to which the rubric indicators represented the intended construct of pedagogical reflection. The analysis used the Content Validity Ratio to assess indicator relevance and Aiken’s V to assess indicator clarity. The results show that most indicators obtained strong expert agreement, indicating that the rubric has adequate content validity for assessing pedagogical reflection in narrative texts.

Table 6. Content Validity of the 4D Reflection Rubric Indicators

Indicator	CVR	Aiken’s V
C1	1.00	0.90
C2	1.00	0.88
C3	1.00	0.92
AR1	1.00	0.87
AR2	0.60	0.82
AR3	0.60	0.80
E1	1.00	0.93
E2	0.60	0.81
E3	1.00	0.90
A1	1.00	0.89
A2	1.00	0.91
A3	1.00	0.88

As presented in Table 6, nine of the twelve indicators achieved a CVR value of 1.00, indicating full agreement among expert validators regarding their relevance. Three indicators,

namely AR2, AR3, and E2, obtained a CVR value of 0.60. Although these values remain acceptable, they suggest that the affective-relational and ethical dimensions required closer refinement because these forms of reflection are often implicit, context-dependent, and more difficult to operationalize consistently.

The Aiken’s V values ranged from 0.80 to 0.93, indicating that all indicators demonstrated satisfactory clarity. The highest clarity score was found in E1, while the lowest was found in AR3. This pattern suggests that the rubric descriptors were generally understandable, but some indicators related to emotional, relational, and ethical reflection required more careful wording to ensure consistent interpretation by raters.

Inter-Rater Reliability of the Rubric

Inter-rater reliability was analyzed to examine whether the rubric could be applied consistently by independent raters. This analysis was particularly important because narrative reflection involves interpretive judgment, especially when reflective meaning is expressed implicitly through the structure of the story rather than through direct explanation. The reliability results are presented in Table 7.

Table 7. Inter-Rater Reliability of the 4D Reflection Rubric

Statistic	Value	Interpretation
ICC(2,1)	0.90	Excellent agreement
95% CI	0.82–0.95	High consistency
F-test	—	Significant, $p < .05$

The ICC(2,1) value of 0.90 indicates excellent agreement among raters. This result suggests that the rubric provides a stable and dependable scoring framework, even though the assessed texts required interpretation of both explicit and implicit reflective elements. The confidence interval of 0.82–0.95 further supports the consistency of the scoring process, indicating that the rubric can be used reliably in assessing teacher-authored narratives.

This finding is important because it demonstrates that a genre-sensitive rubric can still achieve strong reliability when supported by clear indicators, progressive descriptors, and rater calibration. In other words, the interpretive nature of narrative texts does not necessarily reduce scoring consistency when the assessment framework is sufficiently structured.

Distribution of Reflection Scores Across Dimensions

Descriptive statistics were used to examine how pedagogical reflection was distributed across the four dimensions of the rubric. The analysis provides insight into which dimensions appeared more prominently in teacher-authored narratives and which dimensions required further support in teacher professional development.

Table 8. Score Distribution by Reflection Dimension

Dimension	Total Score	Mean
Cognitive reflection	289	9.63
Affective-relational reflection	331	11.03
Ethical reflection	354	11.80
Action-oriented reflection	338	11.27

Table 8 shows that ethical reflection obtained the highest mean score, followed by action-oriented reflection and affective-relational reflection. Cognitive reflection obtained the lowest mean score. This pattern indicates that teachers were more likely to express values, professional responsibility, care, and future-oriented improvement than to explicitly articulate pedagogical reasoning.

The dominance of ethical reflection suggests that teacher-authored narratives often contain strong moral and professional concerns, particularly related to care, responsibility, fairness, and learner well-being. Meanwhile, the relatively lower score for cognitive reflection indicates that teachers may engage in pedagogical reasoning but do not always express it explicitly in narrative form. This finding supports the need for a rubric that can identify both visible and embedded forms of reflection.

The overall scores ranged from 36 to 47, with a mean of 43.50 and a standard deviation of 2.70. This indicates a moderate level of variation in reflective depth among the teacher-authored narratives. The relatively high overall mean also suggests that the selected narratives contained sufficient reflective quality to be assessed using the developed rubric.

Distribution of Reflection Levels

The distribution of reflection levels was analyzed to determine the extent to which teacher-authored narratives moved beyond description toward deeper reflective engagement. The results are presented in Table 9.

Table 9. Distribution of Reflection Levels in Teacher-Authored Narratives

Reflection Level	Frequency	Percentage
Descriptive	3	10.0%
Emerging	8	26.7%
Analytical	12	40.0%
Actionable	7	23.3%

As shown in Table 9, most narratives were categorized at the analytical level, representing 40.0% of the total texts. This indicates that many teachers were able to move beyond simple description and provide meaningful interpretation of classroom experiences. The emerging level accounted for 26.7%, while the actionable level accounted for 23.3%. Only 10.0% of the narratives remained at the descriptive level.

This distribution suggests that teacher reflection in narrative texts should not be simplistically characterized as merely descriptive. Instead, the findings reveal a more nuanced pattern: many teachers were able to analyze classroom experiences, but fewer were able to translate such analysis into explicit future-oriented professional actions. Therefore, the main challenge is not only encouraging teachers to reflect more deeply but also helping them connect reflection with concrete pedagogical improvement.

Differences among raters were most likely to occur at the boundary between emerging and analytical reflection. This occurred particularly when reflective meaning was implied through the narrative sequence rather than explicitly stated. Nevertheless, the strong inter-rater reliability indicates that the rubric remained sufficiently stable in differentiating reflection levels.

Relevance of the Findings to Education Systems and Social Welfare Services

The findings strengthen the relevance of this study to the selected focus of the journal, particularly Education Systems and Social Welfare Services. The journal explicitly welcomes studies related to inclusive education, equitable learning opportunities, workforce professionalism, service quality, institutional capacity building, and evidence-informed educational policy. It also emphasizes research with clear implications for practice, policy, and scalable impact.

In relation to this scope, the present study contributes by offering a validated rubric that can support teacher professional development and improve the quality of reflective assessment in educational institutions. The strong content validity and inter-rater reliability indicate that the rubric is not merely a conceptual framework but also a usable assessment instrument. Its application can help schools, teacher education institutions, and professional learning communities identify the quality of teachers' pedagogical reflection and provide more targeted feedback for professional growth.

The prominence of ethical and action-oriented reflection is also relevant to service quality in education. Teachers who demonstrate ethical awareness and future-oriented reflection are more likely to consider learner well-being, classroom relationships, fairness, and improvement in pedagogical practice. These aspects are closely connected to the broader goal of strengthening educational services that are responsive, humane, and professionally accountable.

Therefore, the results of this study position the 4D Reflection Rubric as an instrument that contributes not only to assessment development but also to institutional capacity building. By enabling more nuanced evaluation of teacher reflection, the rubric can support continuous improvement in teaching practice, professional learning systems, and the broader quality of educational services. This alignment makes the study relevant to the journal's concern with education systems, service quality, and scalable institutional impact.

Discussion

The findings of this study demonstrate that pedagogical reflection in teacher-authored short stories can be understood as a multidimensional construct consisting of cognitive, affective-relational, ethical, and action-oriented reflection. This finding supports the foundational argument that reflection is not merely a cognitive act of reviewing experience, but a professional process through which teachers interpret classroom realities, evaluate decisions, recognize emotional and ethical tensions, and consider future improvement. Daff et al. [27] positioned reflection as an active and disciplined inquiry into experience, while Sherwood [28] emphasized that professional reflection emerges when practitioners confront uncertain and complex situations. The present study extends these perspectives by showing that teachers' reflective competence is represented not only through explicit reasoning but also through emotional awareness, ethical sensitivity, and action-oriented responses embedded in narrative texts.

The identification of four dimensions also aligns with broader studies on reflective teaching, which emphasize that teacher reflection involves multiple layers of professional judgment. Machost and Stains [29] and Gudeta [12] argue that reflective practice requires teachers to examine assumptions, classroom interactions, values, and pedagogical consequences. Similarly, Arendt et al. [30] highlight that reflection should not be reduced to technical evaluation of teaching methods but must include moral, social, and contextual

dimensions. In this regard, the 4D Reflection Rubric provides a more comprehensive framework because it does not assess reflection as a single linear construct. Instead, it recognizes that teachers may reflect through reasoning, relationships, values, and future action in different degrees and forms.

One of the most important findings is that ethical reflection emerged as the most prominent dimension in teacher-authored narratives. This suggests that teachers frequently frame classroom experiences in relation to care, responsibility, fairness, student well-being, and professional values. This finding is consistent with Rambout et al. [31] view that pedagogical action is inherently moral because teachers are constantly required to make judgments about what is appropriate, fair, and meaningful for learners. It also resonates with Marita Cronqvist [32], who argues that teaching is shaped by emotional and ethical practices, not only by instructional techniques. The prominence of ethical reflection in this study indicates that teacher-authored narratives are valuable sources for understanding how teachers negotiate the moral dimensions of classroom life.

However, cognitive reflection was less explicitly articulated than the other dimensions. This finding partially supports earlier studies showing that teachers' reflective writing often remains descriptive or insufficiently analytical [7], [33], [34]. Nevertheless, the present study offers a more nuanced interpretation. The relatively lower cognitive score does not necessarily mean that teachers lack pedagogical reasoning. Rather, it may indicate that teachers do not always express their reasoning explicitly when writing in narrative form. In short stories, pedagogical reasoning may be implied through the sequence of events, the teacher's response to conflict, or the resolution of classroom problems. This distinction is important because it shifts the issue from teachers' reflective incapacity to the alignment between the form of expression and the assessment framework used to evaluate reflection.

The finding that most narratives reached the analytical level, while fewer reached the actionable level, also refines previous assumptions about teacher reflection. Diseth [35] and Vogelsang et al. [36] suggested that many reflective writings tend to remain at descriptive levels. In contrast, the present study shows that teacher-authored narratives can move beyond description and demonstrate meaningful interpretation of classroom experiences. However, the transition from analytical reflection to actionable reflection remains inconsistent. This indicates that the main challenge in reflective practice may not lie only in helping teachers analyze their experiences, but also in supporting them to transform analysis into concrete pedagogical improvement. This finding is important for teacher professional development because reflection becomes more professionally productive when it leads to deliberate changes in future practice.

The narrative nature of the data also provides an important contribution to reflection assessment. Bouizegarene et al. [37] argued that narrative is a mode of meaning-making through which individuals organize experience and construct identity. Coppe et al. [38] similarly emphasized that teacher narratives represent lived experience and professional knowledge. In line with these perspectives, the present study confirms that teacher-authored short stories should not be viewed as informal or secondary forms of reflection. Rather, they can function as rich reflective texts that reveal how teachers interpret pedagogical dilemmas, emotional tensions, relational dynamics, and ethical responsibilities. This supports the argument that assessment tools for reflection should be sensitive to genre and discourse form, especially when reflection is expressed implicitly rather than directly.

The strong content validity and inter-rater reliability of the 4D Reflection Rubric indicate that reflection in narrative texts can be assessed systematically when the instrument is conceptually grounded and operationally clear. This finding is consistent with Kinnear et al. [39] argument that validity concerns the appropriateness of interpretation and use of assessment scores. It also aligns with the Standards for Educational and Psychological Testing, which emphasize that assessment instruments must provide evidence that supports the intended interpretation of scores [40]. In this study, the use of CVR, Aiken's V, and ICC provides evidence that the rubric is not only theoretically relevant but also practically reliable for use by trained raters. This strengthens the validity argument for using the rubric in teacher professional learning contexts.

At the same time, the lower agreement found in several affective-relational and ethical indicators suggests that these dimensions remain more difficult to operationalize than cognitive or action-oriented indicators. This difficulty is understandable because affective and ethical reflection often appears implicitly in narrative texts. For example, emotional awareness may be communicated through tone, dialogue, or character response rather than through direct explanation. Ethical responsibility may also be embedded in the teacher's decision or resolution rather than explicitly stated as a moral principle. This finding reinforces the need for rater calibration and clear descriptors when assessing narrative reflection. It also supports Jonsson et al. [41] view that rubrics are most useful when they provide transparent criteria that guide both assessment and learning.

The findings also have important implications for the use of rubrics as formative tools in teacher professional development. Rubrics are not only scoring instruments but can also function as learning tools that help individuals understand performance expectations, identify strengths, and plan improvement. Karaman [42] show that rubrics can support self-regulated learning when they provide clear criteria and feedback-oriented information. In the context of this study, the 4D Reflection Rubric can help teachers recognize which aspects of their reflection are already strong and which aspects require further development. For example, teachers who demonstrate strong ethical reflection but weak cognitive reflection can be guided to make their pedagogical reasoning more explicit. Similarly, teachers who reach analytical reflection but not actionable reflection can be supported to formulate clearer future teaching strategies.

The results are also relevant to education systems and institutional capacity building. Teacher reflection is closely connected to the quality of educational services because reflective teachers are more capable of evaluating practice, responding to learner diversity, and improving classroom decisions. From this perspective, the rubric developed in this study can support schools, teacher education institutions, and professional learning communities in building more structured systems of reflective practice. Rather than relying on general impressions of teacher reflection, institutions can use the rubric to provide more focused feedback, identify professional development needs, and promote continuous improvement in teaching quality. This contribution is particularly relevant to educational contexts that emphasize student-centered learning, character development, and professional accountability.

The novelty of this study lies in the development of a multidimensional and genre-sensitive rubric for assessing pedagogical reflection in teacher-authored short stories. Previous reflection rubrics have mainly focused on analytical or expository forms of reflective writing, which tend to prioritize explicit reasoning and structured argumentation [43], [44]. In contrast, the 4D Reflection Rubric recognizes that teacher reflection in narrative texts may be expressed

implicitly through events, relationships, emotional responses, ethical tensions, and story resolutions. Therefore, this study contributes a more context-responsive assessment framework that bridges reflective practice, narrative inquiry, and rubric-based assessment.

The implications of this study are both theoretical and practical. Theoretically, the study expands the conceptualization of pedagogical reflection by showing that reflection should be assessed according to the mode of expression through which it appears. Narrative reflection is not a weaker version of analytical reflection; it is a different form of professional meaning-making that requires appropriate assessment criteria. Practically, the rubric provides a validated tool for supporting teacher professional development, reflective writing assessment, and institutional quality improvement. It can be used by teacher educators, school leaders, and professional learning communities to provide more differentiated feedback on teachers' reflective competence. In this way, the study contributes not only to assessment development but also to the broader agenda of improving teacher professionalism, educational service quality, and school-level capacity for continuous improvement.

CONCLUSION

This study concludes that pedagogical reflection in teacher-authored short stories can be systematically assessed through four interconnected dimensions: cognitive, affective-relational, ethical, and action-oriented reflection. The developed 4D Reflection Rubric demonstrated strong content validity, satisfactory indicator clarity, and excellent inter-rater reliability, indicating that it is a valid and reliable instrument for assessing both explicit and implicit forms of teacher reflection in narrative texts. The findings show that ethical and action-oriented reflections were more prominently represented than cognitive reflection, suggesting that teachers tend to express professional values, care, responsibility, and future improvement more clearly than explicit pedagogical reasoning. This study contributes to reflection assessment by offering a multidimensional and genre-sensitive rubric that recognizes narrative writing as a meaningful medium of professional reflection, rather than treating it as a less rigorous form of analytical writing. Practically, the rubric can support teacher professional development, reflective writing assessment, feedback quality, and institutional capacity building in schools, which aligns with the journal's focus on education systems, workforce professionalism, service quality, and scalable institutional impact.

LIMITATIONS

This study has several limitations that should be acknowledged to clarify the scope and interpretation of the findings. First, the study involved a relatively small sample of 30 in-service elementary school teachers who were selected from a national teacher writing community. Although this sample provided relevant narrative data for developing and testing the 4D Reflection Rubric, the participants may have had stronger writing habits and higher engagement with reflective writing than teachers outside such communities. Therefore, the findings should be interpreted cautiously when applied to broader teacher populations or different institutional settings. Second, the study focused only on written narrative texts in the form of teacher-authored short stories. Other forms of reflection, such as oral reflection, collaborative reflection, teaching journals, classroom observation notes, or digital multimodal narratives, were not examined. This limits the extent to which the rubric's applicability can be

generalized across different modes of reflective practice. Third, although the rubric demonstrated strong content validity and excellent inter-rater reliability, the assessment of narrative reflection remains interpretive in nature, particularly when reflective meaning is expressed implicitly through events, dialogue, emotional responses, or story resolutions. This suggests that future use of the rubric should be supported by rater training and calibration to ensure consistent interpretation. Fourth, the study's dissemination stage was limited to refinement and documentation of the validated prototype rather than large-scale implementation in teacher education programs or school-based professional development systems. Further research should therefore involve larger and more diverse samples, different levels of education, and broader institutional contexts to examine the rubric's robustness, scalability, and usefulness for improving teacher professionalism, service quality, and institutional capacity building. This direction is also consistent with the journal's focus on methodologically robust studies with implications for practice, policy, scalable impact, and education systems improvement.

AUTHOR INFORMATION

Corresponding Author

Lilis Ernawati – Universitas Muhammadiyah Malang, (Indonesia);

 orcid.org/0009-0002-2007-3736

Email: ernawatililis433@gmail.com

Authors

Lilis Ernawati – Universitas Muhammadiyah Malang, (Indonesia);

 orcid.org/0009-0002-2007-3736

Ribut Wahyu Eriyanti – Universitas Muhammadiyah Malang, (Indonesia);

 orcid.org/0000-0002-1602-0466

Masduki – Universitas Muhammadiyah Malang, (Indonesia);

 orcid.org/0000-0003-4635-3197

Apriani Riyanti – Universitas Binawan, (Indonesia);

 orcid.org/0000-0002-0140-7326

AUTHOR CONTRIBUTION

All authors contributed to the formulation of the research concept, the development of the study design, and the preparation of the manuscript. L.E. conceptualized the study, designed the research methodology, developed the initial 4D Reflection Rubric, coordinated data collection, conducted the primary data analysis, and led the manuscript writing process. R.W.E. supervised the research design, provided methodological guidance, contributed to the refinement of the theoretical and analytical framework, and critically revised the manuscript for scientific content. M. contributed to the validation of the research instrument, assisted in examining the reliability of rubric-based assessment, and supported the interpretation of the findings. A.R. assisted in data validation, contributed to the refinement of the rubric indicators and descriptors, and supported the final review and editing of the manuscript. All authors have read and approved the final version of the manuscript.

CONFLICT OF INTEREST

"The authors declare no conflict of interest."

DECLARATION OF USE OF AI IN SCIENTIFIC WRITING

The authors used ChatGPT during the preparation of this work to assist in language refinement and structural editing of the manuscript. After utilizing the tool, the authors thoroughly reviewed and revised the content as necessary, and assume full responsibility for the content of this publication.

REFERENCES

- [1] G. A. Koh, H. Askill-Williams, and S. Barr, "Sustaining school improvement initiatives: advice from educational leaders," *Sch. Eff. Sch. Improv.*, vol. 34, no. 3, pp. 298-330, Jul. 2023. <https://doi.org/10.1080/09243453.2023.2190130>.
- [2] M. Faizin, R. Sulaeman, C. Budiyanto, I. Fahmi, and M. Wahid, "Islamic Values-Based HRM and Its Impact on Student Achievement in Islamic Education," *Nidhomul Haq J. Manaj. Pendidik. Islam*, vol. 10, no. 3, pp. 733-748, Dec. 2025. <https://doi.org/10.31538/ndhq.v10i3.243>.
- [3] K. Parveen et al., "The contribution of quality management practices to student performance: Mediated by school culture," *Heliyon*, vol. 10, no. 15, p. e34892, Aug. 2024. <https://doi.org/10.1016/j.heliyon.2024.e34892>.
- [4] Rahmi and S. Rasanjani, "Enhancing teacher quality in Indonesia: The impact of teacher professional development on achieving sustainable development goal 4.c," *Soc. Sci. Humanit. Open*, vol. 12, p. 102123, 2025. <https://doi.org/10.1016/j.ssaho.2025.102123>.
- [5] H. Li, "Reflective Practice for Pre-Service Teachers' Professional Development," *SAGE Open*, vol. 15, no. 3, Jan. 2025. <https://doi.org/10.1177/21582440251363136>.
- [6] Kim, "Impact of Systematic Support in Teacher Education and Professional Development on Training-Teaching Alignment and Instructional Quality," *J. STEM Educ. Res.*, vol. 9, no. 1, pp. 33-58, Jan. 2026. <https://doi.org/10.1007/s41979-025-00154-3>.
- [7] S. Prayogi et al., "Emphasizing reflective processes in scientific inquiry and its impact on preservice science teachers' critical thinking skills," *Soc. Sci. Humanit. Open*, vol. 12, p. 101895, 2025. <https://doi.org/10.1016/j.ssaho.2025.101895>.
- [8] T. Marshall, S. Keville, A. Cain, and J. R. Adler, "Facilitating reflection: a review and synthesis of the factors enabling effective facilitation of reflective practice," *Reflective Pract.*, vol. 23, no. 4, pp. 483-496, Jul. 2022. <https://doi.org/10.1080/14623943.2022.2064444>.
- [9] Y. Kolajo, "Advancing pedagogical excellence through reflective teaching practice and adaptation," *Reflective Pract.*, vol. 26, no. 6, pp. 832-847, Nov. 2025. <https://doi.org/10.1080/14623943.2025.2504143>.
- [10] C. K. Y. Chan and K. K. W. Lee, "Reflection literacy: A multilevel perspective on the challenges of using reflections in higher education through a comprehensive literature review," *Educ. Res. Rev.*, vol. 32, p. 100376, Feb. 2021. <https://doi.org/10.1016/j.edurev.2020.100376>.

- [11] M. Zohrabi and I. Xodabande, "Reflective practice as a tool for teacher education: a comparison between individual and peer reflection of Iranian EFL teachers," *Discov. Educ.*, vol. 3, no. 1, p. 225, Nov. 2024. <https://doi.org/10.1007/s44217-024-00338-w>.
- [12] D. Gudeta, "Professional development through reflective practice: The case of Addis Ababa secondary school EFL in-service teachers," *Cogent Educ.*, vol. 9, no. 1, Dec. 2022. <https://doi.org/10.1080/2331186X.2022.2030076>.
- [13] Z. Chen and R. Chen, "Exploring the Key Influencing Factors on Teachers' Reflective Practice Skill for Sustainable Learning: A Mixed Methods Study," *Int. J. Environ. Res. Public Health*, vol. 19, no. 18, p. 11630, Sep. 2022. <https://doi.org/10.3390/ijerph191811630>.
- [14] D. P. Babić and K. M. Zdravković, "Empowering Education Through Collaboration: School as a Learning Community," 2026, pp. 177-193. https://doi.org/10.1007/978-3-032-19894-5_10.
- [15] J. Trzebiński, J. Z. Czarnecka, and M. Cabański, "The impact of the narrative mindset on effectivity in social problem solving," *PLoS One*, vol. 16, no. 7, p. e0253729, Jul. 2021. <https://doi.org/10.1371/journal.pone.0253729>.
- [16] S. Owiti, T. Bersier, and D. Hauw, "Individual differences in professional sport narrative experience during basketball players club mutation," *Heliyon*, vol. 7, no. 9, p. e08015, Sep. 2021. <https://doi.org/10.1016/j.heliyon.2021.e08015>.
- [17] N. N. Makarem et al., "Grading reflective essays: the reliability of a newly developed tool- GRE-9," *BMC Med. Educ.*, vol. 20, no. 1, p. 331, Dec. 2020. <https://doi.org/10.1186/s12909-020-02213-2>.
- [18] H. Shafiee Rad, "Reinforcing L2 reading comprehension through artificial intelligence intervention: refining engagement to foster self-regulated learning," *Smart Learn. Environ.*, vol. 12, no. 1, p. 23, Mar. 2025. <https://doi.org/10.1186/s40561-025-00377-2>.
- [19] N. T. Hang, "Vietnamese EFL Teachers' Readiness and Challenges in Integrating AI-Generated Feedback," *Lang. Technol. Soc. Media*, 2025. <https://doi.org/10.70211/ltsm.3026-7196.315>.
- [20] N. F. Asli, M. E. E. Mohd Matore, and M. Md Yunus, "Construct validity of primary trait writing rubrics based on assessment use argument (AUA) validation framework," *Heliyon*, vol. 10, no. 22, p. e40053, Nov. 2024. <https://doi.org/10.1016/j.heliyon.2024.e40053>.
- [21] M. C. Tavares and M. Vaz, "Rethinking Performance Evaluation: Strategic Alignment in the Service Sector Through a Case-Based Framework," *Adm. Sci.*, vol. 15, no. 10, p. 390, Oct. 2025. <https://doi.org/10.3390/admsci15100390>.
- [22] M. Vuorre and J. Metcalfe, "Measures of relative metacognitive accuracy are confounded with task performance in tasks that permit guessing," *Metacognition Learn.*, vol. 17, no. 2, pp. 269-291, Aug. 2022. <https://doi.org/10.1007/s11409-020-09257-1>.
- [23] Hendriyanto, D. Suryadi, S. Sahara, D. Fardian, I. Pauji, and L. H. Muhaimin, "From tools to thought partners: Optimizing technology as extended cognition for innovative didactic design," *AIP Conf. Proc.*, vol. 3220, no. 1, pp. 1-9, Oct. 2024. <https://doi.org/10.1063/5.0234677>.

- [24] S. P. Phokoye et al., "Exploring the Adoption of Robotics in Teaching and Learning in Higher Education Institutions," *Informatics*, vol. 11, no. 4, p. 91, Nov. 2024. <https://doi.org/10.3390/informatics11040091>.
- [25] M. Farrús, "Automatic Speech Recognition in L2 Learning: A Review Based on PRISMA Methodology," *Languages*, vol. 8, no. 4, p. 242, Oct. 2023. <https://doi.org/10.3390/languages8040242>.
- [26] S. Thiagarajan, D. S. Semmel, and M. I. Semmel, *Instructional development for training teachers of exceptional children*. Minnesota: Indiana University, 1974.
- [27] L. Daff, C. Tame, and J. Sands, "A course design approach that encourages reflective practice habits," *Int. J. Manag. Educ.*, vol. 22, no. 2, p. 100990, Jul. 2024. <https://doi.org/10.1016/j.ijme.2024.100990>.
- [28] G. Sherwood, "Reflective practice and knowledge development: Transforming research for a practice-based discipline," *Int. J. Nurs. Sci.*, vol. 11, no. 4, pp. 399-404, Sep. 2024. <https://doi.org/10.1016/j.ijnss.2024.08.002>.
- [29] H. Machost and M. Stains, "Reflective Practices in Education: A Primer for Practitioners," *CBE-Life Sci. Educ.*, vol. 22, no. 2, Jun. 2023. <https://doi.org/10.1187/cbe.22-07-0148>.
- [30] K. Arendt, L. Stark, A. Friedrich, R. Brünken, and R. Stark, "Quality of Reflections on Teaching: Approaches to Its Measurement and Low-Threshold Promotion," *Educ. Sci.*, vol. 15, no. 7, p. 884, Jul. 2025. <https://doi.org/10.3390/educsci15070884>.
- [31] F. Rombout, J. A. Schuitema, and M. L. L. Volman, "Teaching strategies for value-loaded critical thinking in philosophy classroom dialogues," *Think. Ski. Creat.*, vol. 43, p. 100991, Mar. 2022. <https://doi.org/10.1016/j.tsc.2021.100991>.
- [32] M. Cronqvist, "Teachers' ethical responsibility in teaching; to guide the children about right and wrong," *Scand. J. Educ. Res.*, vol. 69, no. 4, pp. 857-870, Jun. 2025. <https://doi.org/10.1080/00313831.2024.2360901>.
- [33] M. Gläser-Zikuda, C. Zhang, F. Hofmann, L. Plöbl, L. Pösse, and M. Artmann, "Mixed methods research on reflective writing in teacher education," *Front. Psychol.*, vol. 15, Oct. 2024. <https://doi.org/10.3389/fpsyg.2024.1394641>.
- [34] B. A. Dean et al., "Teaching reflection in higher education curriculum: evaluation of student learning through a pre- and post-intervention," *Reflective Pract.*, pp. 1-13, Apr. 2026. <https://doi.org/10.1080/14623943.2026.2661198>.
- [35] Å. Diseth, "Reflection notes as a tool for professional development of pre-service psychology teachers," *Reflective Pract.*, vol. 26, no. 6, pp. 821-831, Nov. 2025. <https://doi.org/10.1080/14623943.2025.2504139>.
- [36] C. Vogelsang, D. Scholl, J. Meier, and S. Küth, "Thoughts Are Free-Differences Between Unstructured and Structured Reflections of Teachers with Different Levels of Expertise," *Educ. Sci.*, vol. 15, no. 7, p. 820, Jun. 2025. <https://doi.org/10.3390/educsci15070820>.
- [37] N. Bouizegarene, M. J. D. Ramstead, A. Constant, K. J. Friston, and L. J. Kirmayer, "Narrative as active inference: an integrative account of cognitive and social functions in adaptation," *Front. Psychol.*, vol. 15, Jun. 2024. <https://doi.org/10.3389/fpsyg.2024.1345480>.
- [38] T. Coppe, M. Parmentier, G. Kelchtermans, I. Raemdonck, V. März, and S. Colognesi, "Beyond traditional narratives about teacher professional development: A critical

- perspective on teachers' working life," *Teach. Teach. Educ.*, vol. 139, p. 104436, Mar. 2024. <https://doi.org/10.1016/j.tate.2023.104436>.
- [39] B. Kinnear, C. St-Onge, D. J. Schumacher, M. Marceau, and T. Naidu, "Validity in the Next Era of Assessment: Consequences, Social Impact, and Equity," *Perspect. Med. Educ.*, vol. 13, no. 1, pp. 452-459, Sep. 2024. <https://doi.org/10.5334/pme.1150>.
- [40] G. H. Poll, M. Mayes, and J. Petru, "Evidence to support test interpretation: Evaluating the content validity of a new item set for the Transition Pragmatics Interview," *Int. J. Speech. Lang. Pathol.*, vol. 26, no. 6, pp. 809-819, Nov. 2024. <https://doi.org/10.1080/17549507.2023.2287424>.
- [41] Jonsson, E. Panadero, L. Pinedo, and B. Fernández-Castilla, "Using rubrics for formative purposes: identifying factors that may affect the success of rubric implementations," *Assess. Educ. Princ. Policy Pract.*, vol. 32, no. 2, pp. 192-211, Mar. 2025. <https://doi.org/10.1080/0969594X.2025.2486947>.
- [42] P. Karaman, "Effects of using rubrics in self-assessment with instructor feedback on pre-service teachers' academic performance, self-regulated learning and perceptions of self-assessment," *Eur. J. Psychol. Educ.*, vol. 39, no. 3, pp. 2551-2574, Sep. 2024. <https://doi.org/10.1007/s10212-024-00867-w>.
- [43] H. T. Gebremariam, M. A. Tessema, M. T. Kahsay, Z. A. Mulugeta, T. W. Ambaye, and A. T. Mulatu, "Using integrated alternative assessment to explore essay writing skills: An insight into authentic instructional dashboards," *Soc. Sci. Humanit. Open*, vol. 11, p. 101571, 2025. <https://doi.org/10.1016/j.ssaho.2025.101571>.
- [44] K. Kager, A. Jurczok, S. Bolli, and M. Vock, "'We were thinking too much like adults': Examining the development of teachers' critical and collaborative reflection in lesson study discussions," *Teach. Teach. Educ.*, vol. 113, p. 103683, May 2022. <https://doi.org/10.1016/j.tate.2022.103683>.