2026 SCIENCELINE
Journal of Educational and Management Studies
Volume 16, Issue 1: 01-09; March 25, 2026

ISSN: 2322-4770

DOI: https://dx.doi.org/10.54203/jems.2026.1

Refining instructional materials evaluation: adoption of a rating scale in the institutional evaluation tool

Jesus Parena Santillan[1], and Dolores C. Volante [2]

[1] *Faculty Member, Camarines Norte State College, Philippines*

[2] *Vice President for Academic Affairs/Faculty Member, Camarines Norte State College, Philippines*

Email: jesussantillan@cnsc.edu.ph; doloresvolante@cnsc.edu.ph

ABSTRACT

Effective evaluation of instructional materials (IM) is essential for ensuring educational quality in higher education. This institutional case study examines the refinement of Camarines Norte State College's existing policy (BOT Res. No. 60, s. 2022), which previously lacked a defined rating scale, resulting in inconsistencies and subjectivity in IM assessment. The study aimed to develop a transferable policy refinement model by introducing a structured rating score evaluation scale that enhances clarity, standardization, transparency, and usability while preserving policy integrity. Data were collected through focus group discussions with faculty and evaluators and document analysis, and analyzed using thematic analysis. Key themes identified include policy alignment, standardization, accountability, efficiency, and participatory governance. The resulting framework offers a scalable mechanism for higher education institutions to bridge qualitative evaluation with quantitative accountability, supporting evidence-based decision-making and continuous improvement in instructional material assessment.

Original Article

PII: S2322477026000001-16

Rec. 08 November, 2025
Acc. 17 March, 2026
Pub. 25 March, 2026

Keywords

**Instructional Materials Evaluation,
Rating Score Scale,
Higher Education Policy**

INTRODUCTION

In the contemporary institutional landscape, commitment to ethical standards and accountability has emerged as a crucial determinant of organizational credibility and long-term sustainability (AlHares, 2025; Stavropoulou et al., 2024). While these principles have long been emphasized in the corporate sector through initiatives such as Corporate Social Responsibility (CSR), they have become equally vital within higher education institutions. Colleges and universities are increasingly expected to uphold transparency, fairness, and ethical governance to sustain public trust and institutional integrity (Delgado-Alemany et al., 2021; Gichuru, 2023). Similar to corporations where 76% of consumers report discontinuing engagement with companies that neglect social and ethical responsibility (PwC, 2021) academic institutions must likewise demonstrate accountability and responsiveness to stakeholders. Aligning ethical standards with institutional practices fosters a culture of trust, integrity, and quality assurance essential for sustainable institutional development.

In higher education, ethical accountability is operationalized through mechanisms that ensure the quality and effectiveness of instruction (Harvey & Williams, 2010). Among these mechanisms, evaluation tools serve as vital instruments for assessing the pedagogical soundness, relevance, and institutional alignment of instructional materials (Biggs & Tang, 2011). These tools reinforce transparency and objectivity in academic assessment, enabling institutions to uphold teaching excellence and promote continuous improvement (Bertolin, 2021). However, the effectiveness of such instruments depends largely on the clarity, consistency, and interpretability of their rating criteria.

Previous studies emphasize the importance of expert judgment, qualitative review, and professional discretion in evaluating instructional materials (Harvey & Williams, 2010; Biggs & Tang, 2011). Qualitative evaluation allows evaluators to consider contextual relevance, pedagogical appropriateness, and instructional creativity. Nevertheless, this reliance on subjective judgment also introduces variability in scoring and interpretation, which may compromise consistency and comparability across evaluators (Jonsson & Svingby, 2007). In contrast, recent quality assurance frameworks increasingly advocate for quantifiable, standardized, and data-driven evaluation systems to strengthen institutional

To cite this paper: Santillan J.P., and Volante D.C. (2026). Refining instructional materials evaluation: adoption of a rating scale in the institutional evaluation tool. *J. Educ. Manage. Stud.*, 16(1):01-09. DOI: https://dx.doi.org/10.54203/jems.2026.1

accountability and transparency (Panadero & Jonsson, 2013; Bertolin, 2021).

This situation reveals a persistent tension in the literature between qualitative, expert-based assessment and the growing demand for quantitative accountability in higher education evaluation systems. While qualitative approaches promote depth and contextual sensitivity, they often lack standardized mechanisms for score interpretation. Conversely, purely quantitative systems may enhance comparability but risk oversimplifying complex instructional qualities (Sadler, 2009). Despite extensive research on assessment tools and rubric design, limited studies have focused on how institutions can systematically integrate qualitative judgment with standardized rating scales within existing policy frameworks.

At Camarines Norte State College (CNSC), this tension is reflected in the long-standing use of an approved Institutional Instructional Materials Evaluation Tool that relies heavily on professional judgment without a clearly defined rating score scale. Prior to this study, the absence of a standardized scoring framework resulted in inconsistencies in how evaluators interpreted and assigned scores, thereby weakening inter-rater reliability and limiting the comparability of evaluation outcomes. Similar challenges have been reported in previous studies highlighting the need for clearly articulated scoring systems to ensure fairness and objectivity in educational assessment (Jonsson & Svingby, 2007; Panadero & Jonsson, 2013).

The present study addresses this institutional and scholarly gap through an institutional case study approach that examines and refines the existing evaluation system of CNSC. Specifically, it responds to a critical limitation in the Revised Policies and Guidelines on Instructional Material Development under BOT Resolution No. 60, s. 2022, which lacks a defined rating score evaluation scale. Rather than revising established policies, this study proposes a policy refinement model through the adoption of a standardized rating score evaluation scale designed to complement the approved evaluation tool.

The proposed scale introduces a transparent and quantifiable method for interpreting evaluation results. By translating numerical scores into meaningful qualitative descriptors, it bridges the gap between professional judgment and institutional accountability (Brown & White, 2021). This hybrid approach strengthens evaluation reliability while preserving the pedagogical sensitivity of expert review. Importantly, this refinement functions as a policy enhancement mechanism that improves evaluation efficiency, reliability, and fairness without altering existing regulations (Green, 2022). As such, the framework developed in this study may serve as a transferable model for other higher education institutions experiencing similar tensions between qualitative assessment practices and accountability requirements.

Moreover, this initiative responds to the institution's immediate operational need to assess a backlog of previously submitted instructional materials. The adoption of a standardized rating score system enables CNSC to conduct evaluations more systematically and equitably, in accordance with established quality assurance principles. The approach also aligns with recognized grading and assessment practices, facilitating institutional acceptance and practical implementation.

Hence, this study was undertaken to refine and institutionalize a clear and consistent rating score evaluation scale aligned with the approved Institutional Instructional Materials Evaluation Tool of CNSC. It represents a policy refinement model aimed at strengthening transparency, fairness, and comparability in instructional material evaluation while reinforcing the institution's internal quality assurance framework (Harvey & Green, 1993; Sadler, 2009). Through participatory feedback from faculty members, evaluators, and academic administrators, the proposed scale was systematically refined and validated. The resulting framework promotes a more reliable, transparent, and equitable approach to instructional material assessment and contributes to CNSC's pursuit of academic excellence and continuous quality improvement. Furthermore, the study offers a practical reference for other higher education institutions seeking to enhance their evaluation systems through evidence-based policy refinement.

Conceptual framework

This study was anchored on the principles of institutional accountability and continuous quality assurance in higher education and is framed as an institutional case study and policy refinement model. It builds on the Revised Policies and Guidelines on Instructional Material (IM) Development under BOT Resolution No. 60, s. 2022, which provides the institutional foundation for evaluating instructional materials at Camarines Norte State College (CNSC).

However, while the policy outlines key components for evaluating instructional materials, it lacks a defined rating score evaluation scale necessary to ensure uniformity and objectivity in the scoring process. To address this gap, the present study refined and institutionalized a rating score evaluation scale designed to complement, rather than alter, the

approved IM Evaluation Tool. This refinement model may serve as a transferable framework for other higher education institutions facing similar evaluation gaps.

The conceptual framework is guided by Sadler's (2009) theory of formative assessment and Harvey and Green's (1993) quality dimensions in higher education, emphasizing that consistent evaluation mechanisms and transparent standards are critical to ensuring fairness, credibility, and continuous improvement in academic processes. Within this context, the study adopted a qualitative-developmental approach, engaging key stakeholders—faculty members, instructional material evaluators, and administrators—through focus group discussions (FGDs) to refine the proposed scale, as illustrated in Figure 1.

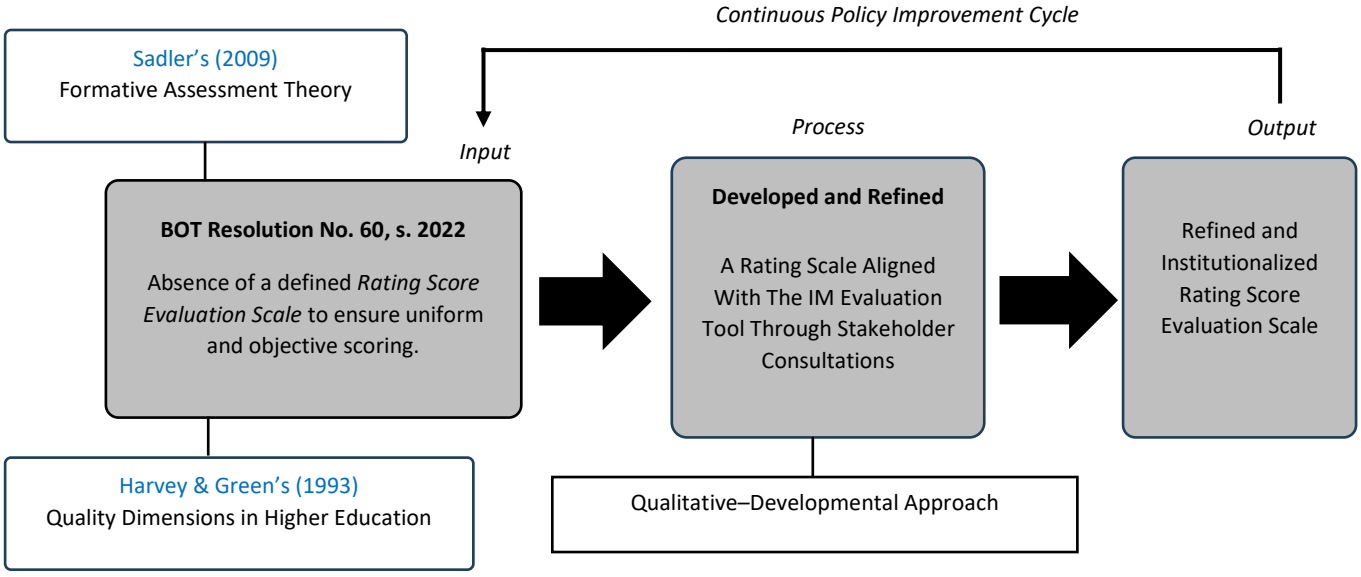

Figure 1. Paradigm of the Study

Through this participatory process, stakeholder insights provided the empirical basis for enhancing the clarity, applicability, and fairness of the rating score descriptors and intervals. The refined scale, once integrated with the existing IM Evaluation Tool, strengthens CNSC's internal quality assurance system by:

(1) Ensuring transparency and consistency in instructional material evaluation;

(2) Promoting objectivity and inter-rater reliability; and

(3) Aligning institutional practices with recognized standards of educational quality management.

Ultimately, the framework illustrates a cyclical process in which institutional policy (BOT Resolution No. 60, s. 2022) informs the refinement of the evaluation tool, stakeholder validation ensures contextual alignment, and the adopted scale feeds back into enhanced institutional quality assurance—thereby sustaining academic excellence, public trust, and ethical accountability.

METHODOLOGY

Research Design

This study employed a qualitative developmental research design, which was appropriate for refining and institutionalizing an evaluation scale based on stakeholder feedback. The developmental approach focused on improving an existing institutional framework, the Revised Policies and Guidelines on Instructional Material (IM) Development under BOT Resolution No. 60, s. 2022 by introducing a complementary rating score evaluation scale.

This design enabled the systematic integration of expert insights and institutional perspectives to ensure the contextual validity, applicability, and potential transferability of the refined scale (Richey & Klein, 2014). The qualitative component emphasized participatory engagement and thematic refinement, consistent with formative research design frameworks that support the iterative development of educational instruments and policies (Ruiz-Primo & Li, 2017; Patton, 2015).

Participants

The participants of the study consisted of faculty members and instructional material (IM) evaluators who participated in the first round of focus group discussions, and members of the Academic Council who took part in the second round of review and presentation at Camarines Norte State College (CNSC). These stakeholders were purposively selected because of their direct involvement in the institutional processes of preparing, reviewing, and approving instructional materials.

Purposive sampling was employed to ensure that only individuals with relevant expertise and experience contributed to the refinement of the evaluation scale. A total of 12 participants took part in the initial focus group discussion (FGD), comprising six faculty members, four IM evaluators, and two administrators. Their diverse institutional roles ensured balanced representation of perspectives on instructional material development and assessment.

Data gathering procedure

Data were gathered through two stages of focus group discussions (FGDs) and document analysis. Initially, the researcher examined the Revised Policies and Guidelines on Instructional Material Development (BOT Resolution No. 60, s. 2022) and the existing IM Evaluation Tool to identify structural gaps, particularly the absence of a defined rating scale. This review served as the foundation for developing the initial version of the proposed rating score evaluation scale.

The first FGD sought participants' feedback on the clarity, applicability, and relevance of the initial draft. Participants' comments and suggestions were documented and subjected to thematic analysis to identify recurring issues and recommendations. Based on these insights, the researcher revised the draft scale and presented the updated version for validation through member checking to ensure that revisions accurately reflected participants' perspectives.

After the initial revision, a second FGD was conducted to present the refined version of the scale. During this phase, members of the Academic Council reviewed the proposed policy and guidelines and provided qualitative feedback for further improvement. The results informed the final revision of the rating score evaluation scale, which was subsequently endorsed for institutional presentation and adoption. Member checking was again conducted prior to submission to the Board of Trustees for final approval.

Research instrument

The primary instrument of the study was the proposed Rating Score Evaluation Scale for Instructional Materials, which was developed and refined through a series of stakeholder consultations. The initial version consisted of five rating categories with corresponding qualitative descriptors designed to interpret numerical results from the existing IM Evaluation Tool.

During the FGDs, participants evaluated the clarity, relevance, and fairness of each rating category. Their feedback guided revisions in phrasing, score intervals, and descriptive labels. The final instrument was the refined and institutionalized rating scale, integrated as a supplementary component of the existing evaluation tool.

Expert validation was conducted by the Academic Council Committee on Quality Assurance to establish content validity and ensure alignment with institutional policy standards.

Data analysis

The qualitative data gathered from the focus group discussions were transcribed verbatim and analyzed using Braun and Clarke's (2006) six-phase thematic analysis framework. This approach was employed to systematically identify patterns and shared meanings in participants' responses related to the refinement of the evaluation scale.

The analysis followed six iterative phases: 1) familiarization through repeated reading of transcripts; 2) initial code generation using open coding; 3) searching for themes by clustering related codes; 4) reviewing themes for coherence and consistency; 5) defining and naming final themes; and 6) producing analytic narratives for interpretation.

To enhance analytic rigor, coding was conducted in two cycles. The first cycle involved inductive generation of preliminary codes, while the second cycle focused on refining, merging, and reorganizing codes based on thematic relationships and theoretical alignment. Selected transcripts and preliminary codes were reviewed by a peer researcher to minimize subjective interpretation. Discrepancies were resolved through discussion and consensus-building.

The finalized themes—clarity, objectivity, fairness, and usability—were compared with established frameworks in educational evaluation and quality assurance, particularly those emphasizing transparency, validity, and reliability in rubric construction (Sirimanna et al., 2022). These comparisons guided the refinement of qualitative descriptors and score intervals.

To ensure trustworthiness, the study employed strategies for credibility, dependability, confirmability, and transferability. Member checking was conducted by returning thematic summaries to participants for verification. Triangulation was achieved through the integration of FGD data and document analysis. An audit trail of analytic decisions was maintained to support methodological rigor and minimize potential researcher bias.

Ethical considerations

Prior to data collection, ethical clearance was obtained from the institutional Research Ethics Committee. Participants were fully informed of the study's objectives, procedures, and voluntary nature. Written informed consent was secured, emphasizing confidentiality, anonymity, and the right to withdraw without penalty.

All data collected were treated with strict confidentiality. No identifying information was included in the transcriptions, analyses, or published reports. Data storage complied with institutional ethical guidelines and the Data Privacy Act of 2012 (Republic Act No. 10173) of the Philippines. Ethical standards were upheld throughout all phases of the study to protect participants' dignity and data security.

RESULTS AND DISCUSSION

Refined policy and guidelines on the rating score evaluation scale

The thematic analysis of the finalized policy entitled "*Adoption of the Rating Score Evaluation Scale for the Existing Approved Institutional Instructional Materials Evaluation Tool of Camarines Norte State College*" (see Appendix A) revealed its significant contribution to strengthening institutional assessment mechanisms. The integration of a standardized rating score evaluation scale directly addressed the gap identified in Board of Trustees (BOT) Resolution No. 60, s. 2022, which previously lacked a clearly defined numerical evaluation framework.

While the original policy emphasized qualitative judgment and expert review, it provided limited guidance for consistent numerical interpretation. This reflects a broader tension in higher education assessment between qualitative professional discretion and the growing demand for quantitative accountability. The revised policy responded to this tension by establishing a transparent and quantifiable system that supports uniform interpretation of instructional material (IM) evaluation results while preserving expert judgment.

The refined policy aligns with institutional quality assurance objectives by integrating numerical ranges with qualitative descriptors, thereby strengthening both procedural consistency and evaluative credibility. Table 1 presents the thematic analysis of the refined policy and guidelines on the adoption of the rating score evaluation scale for instructional materials.

Findings indicate that the refined policy was firmly anchored in BOT Resolution No. 60, s. 2022. Rather than replacing the existing framework, the scale complemented and strengthened institutional regulations. This alignment enhanced institutional legitimacy and facilitated administrative acceptance. Consistent with higher education policy evaluation literature, reforms grounded in existing governance structures are more likely to achieve sustainability and compliance.

The introduction of defined numerical ranges (e.g., 1.00–1.79 "Needs Major Revision" to 4.20–5.00 "Outstanding") provided a standardized benchmark for interpreting evaluation results. The explicit computation formula—dividing the total achieved score by the number of indicators—further strengthened procedural clarity and reproducibility.

This standardized structure addressed long-standing concerns regarding inter-rater variability and subjective interpretation. Previous studies emphasize that while qualitative judgment remains essential, the absence of clear scoring criteria may compromise consistency. The present findings demonstrate that integrating quantitative benchmarks with descriptive indicators enhances both reliability and interpretive depth (Jonsson & Svingby, 2007; Brookhart, 2018).

The combined use of numerical scores and qualitative descriptors ("Outstanding," "Very Good," "Good," "Fair," and "Needs Major Revision") strengthened accountability by making evaluative decisions transparent and traceable. This dual system facilitated evidence-based decisions regarding approval, revision, or rejection of instructional materials.

Moreover, the scale functioned as a governance tool that supported structured institutional decision-making. Clear evaluative thresholds reduced ambiguity in administrative deliberations and promoted consistency across academic departments. Such governance clarity contributes to public trust in institutional evaluation processes and reinforces institutional credibility.

Table 1. Thematic analysis of the refined policy and guidelines on the adoption of the rating score evaluation scale for instructional materials

Theme	Sub-theme / Indicator	Data Basis / Verbatim Evidence	Interpretation / Implication
Policy Alignment & Institutional Legitimacy	Anchored in BOT Res. No. 60, s. 2022; maintains existing framework	"This proposal seeks to address a critical gap of the approved revised policies and guidelines on Instructional Material (IM) Development, BOT Res. No. 60, s. 2022."	The policy complements rather than replaces existing regulations, enhancing institutional legitimacy and sustainability. Ensures compliance with governance structures.
Standardization & Objectivity in Evaluation	Defined rating scores and ranges; numeric + qualitative descriptors	"It lacks a defined rating scale — a fundamental feature necessary for ensuring clarity and consistency in the assessment process."	Introducing a standardized scoring system addresses inter-rater reliability, improves consistency, and ensures fairness in evaluation.
Transparency & Accountability	Clear interpretation of scores; decision-classification	"By interpreting raw scores using the rating scale ranges with corresponding qualitative descriptors, this system adds a layer of granularity and objectivity."	Numeric + descriptive system allows evaluative decisions to be traceable, promoting accountability and institutional credibility.
Efficiency & Practical Implementation	Reduces IM evaluation backlog; applies without overhauling policy	"By adopting this temporary rating score system, the institution can efficiently address evaluation backlogs..."	Streamlines evaluation procedures, reduces administrative burden, and expedites decision-making.
Usability & Evaluator Adaptability	Aligns with standard academic grading; easy to use	"The adoption of a rating scoring system aligns with the grading practices commonly used in educational institutions."	Familiar scoring system reduces training needs, improves evaluator confidence, and ensures consistent application across departments.
Quality Assurance & Continuous Improvement	Integration with existing rubrics; supports targeted feedback	"By employing rating score ranges, evaluators can more effectively identify specific strengths and areas for improvement in instructional materials."	Facilitates iterative improvement of instructional materials, supporting a culture of continuous quality enhancement.
Decision Support & Governance Clarity	Structured decision-making matrix; clear approval/revision/rejection criteria	"It provides clear criteria for determining whether an IM should be approved, recommended with revisions, or rejected."	Promotes evidence-based governance and transparent decision-making; enhances administrative clarity and public trust.
Stakeholder Engagement & Institutional Ownership	Iterative revisions through FGDs and Academic Council meetings; participatory process	"Presented: Institutional IM Committee ... Presented: Academic Council Meeting."	Participatory design increases stakeholder buy-in, ensuring the instrument is valid, reliable, and institutionally owned.

The analysis yielded five dominant themes: (1) policy alignment and legitimacy, (2) standardization and objectivity, (3) transparency and accountability, (4) usability and accessibility, and (5) participatory governance.

The findings also highlighted the practical advantages of the refined scale. Its alignment with familiar academic grading practices enhanced usability and reduced the need for extensive retraining. As a result, evaluators demonstrated greater adaptability and confidence in applying the instrument.

Furthermore, the scale addressed operational challenges such as evaluation backlogs by streamlining assessment procedures without requiring major policy revisions. This efficiency reflects best practices in educational policy implementation, where minor but strategic refinements can yield substantial administrative benefits (OECD, 2020).

The participatory nature of the policy refinement process emerged as a central theme. The involvement of the Institutional IM Committee and the Academic

Council ensured that diverse stakeholder perspectives were incorporated. This collaborative approach strengthened institutional ownership and enhanced the acceptability of the revised instrument.

Consistent with collaborative policy development frameworks, stakeholder engagement contributed to both the technical validity and the social legitimacy of the scale (McNamara et al., 2015). Such participatory mechanisms are essential for sustaining long-term institutional reforms.

Beyond procedural improvement, the refined evaluation scale has broader implications for instructional quality and academic development. By providing evaluators with precise diagnostic feedback, the scale facilitates targeted revision of instructional materials. This, in turn, supports the development of higher-quality learning resources and contributes to improved student learning outcomes.

Standardized and transparent evaluation practices also reinforce a culture of research-based material development, encouraging faculty members to engage in systematic instructional design and continuous improvement. Over time, this may strengthen the institution's research culture and commitment to evidence-based pedagogy.

Despite its strengths, the refined policy does not entirely eliminate evaluator subjectivity. Although numerical ranges and descriptors enhance consistency, interpretation may still vary across evaluators and departments. Continuous training and calibration remain necessary to ensure uniform application of the scale.

Future studies may examine the long-term reliability of the instrument through quantitative inter-rater reliability testing and longitudinal monitoring of evaluation outcomes.

In summary, the thematic analysis demonstrates that the adoption of the Rating Score Evaluation Scale represents more than a procedural adjustment. It constitutes a systematic advancement in institutional quality assurance and policy governance. By bridging qualitative evaluation and quantitative accountability, the refined policy promotes fairness, objectivity, transparency, and sustainability in instructional material assessment.

Its integration into BOT Resolution No. 60, s. 2022 reflects an institutional commitment to data-informed decision-making, enhanced teaching quality, and continuous academic improvement. As a policy refinement model, the framework may also serve as a reference for other higher education institutions seeking to strengthen their instructional evaluation systems.

CONCLUSION

The refined policy on the Adoption of the Rating Score Evaluation Scale for the Existing Approved Institutional Instructional Materials Evaluation Tool effectively addressed the gap in BOT Resolution No. 60, s. 2022, by providing a standardized quantitative framework for evaluating instructional materials. The integration of defined numerical ranges with qualitative descriptors enhanced objectivity, transparency, and fairness, ensuring consistent and reproducible evaluative judgments. Anchored within institutional policy, the scale improved efficiency, inter-rater reliability, and participatory governance, reflecting a collaborative and evidence-based approach to institutional quality assurance. Overall, this policy represents a significant advancement in data-driven decision-making and continuous improvement in instructional material assessment at Camarines Norte State College.

Recommendations

1. The rating score evaluation scale should be formally adopted and implemented across all academic units, accompanied by structured evaluator training to ensure uniform application.
2. Regular monitoring and periodic review of the scale's effectiveness are recommended to maintain its relevance and reliability.
3. Integration of the scale into a digital evaluation system is encouraged to enhance efficiency, transparency, and accessibility for evaluators and administrators.
4. Future studies may consider assessing longitudinal impacts on instructional quality and inter-rater consistency to further validate the instrument's effectiveness.

DECLARATIONS

Corresponding author

Correspondence and requests for materials should be addressed to Jesus Parena Santillan; E-mail: jesussantillan@cnsc.edu.ph; ORCID: https://orcid.org/0000-0002-6976-7790, and Dolores C. Volante; E-mail: doloresvolante@cnsc.edu.ph; ORCID: https://orcid.org/0000-0002-2236-5655.

Data availability

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

Supplementary Information

There is no online version contains supplementary material available

Authors' contribution

J.P. Santillan contributed to the overall research, including data analysis and manuscript writing. D.C. Volantes contributed by assisting in the approval of the research output and providing comments to enhance the refined evaluation tool developed in the study.

Acknowledgements

The authors thank Camarines Norte State College for providing the resources necessary for this research and for supporting the adoption of the developed evaluation instrument for instructional materials. The authors also extend their gratitude to all experts involved in reviewing, polishing, and enhancing the proposed evaluation instrument.

Funding support

The authors declare that no funds, grants, or other support were received during the preparation or publication of this manuscript.

Competing interests

The authors declare no competing interests in this research and publication.

REFERENCES

AlHares, A. (2025). Ethical leadership and its impact on corporate sustainability and financial performance: The role of alignment with the Sustainable Development Goals. *Sustainability, 17*(15), 6682. https://doi.org/10.3390/su17156682

Delgado-Alemany, R., Revilla-Camacho, M. Á., & Blanco-González, A. (2021). Is a university committed to ethics perceived as an honest, appropriate and properly managed organization? *Journal of Management and Business Education, 24*(1). https://doi.org/10.35564/jmbe.2021.0001

Gichuru, E. (2023). Ethical standards in higher education. *Kiu Journal of Education, 3*(2), 98-114. https://doi.org/10.59568/KJED-2023-3-2-11

PwC. (2021). *Beyond compliance: Consumers and employees want business to do more on ESG.* https://www.pwc.com/us/en/services/consulting/library/consumer-intelligence-series/consumer-and-employee-esg-expectations.html

Stavropoulou, A., Ooms, A., Rovithis, M., & Stroumpouki, T. (2024). Key concepts of ethical leadership: A review of the literature. *World Journal of Advanced Research and Reviews, 21*(3), 2227-2232. https://doi.org/10.30574/wjarr.2024.21.3.0913

Bertolin, J. C. G. (2021). Higher education quality assurance systems: International models and challenges for developing countries. *Higher Education Quarterly, 75*(3), 475–490. https://doi.org/10.1111/hequ.12285

Biggs, J., & Tang, C. (2011). *Teaching for quality learning at university* (4th ed.). McGraw-Hill Education.

Harvey, L., & Green, D. (1993). Defining quality. *Assessment & Evaluation in Higher Education, 18*(1), 9–34. https://doi.org/10.1080/0260293930180102

Harvey, L., & Williams, J. (2010). Fifteen years of quality in higher education. *Quality in Higher Education, 16*(1), 3–36. https://doi.org/10.1080/13538321003679457

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

Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review, 9*, 129–144. https://doi.org/10.1016/j.edurev.2013.01.002

Sadler, D. R. (2009). Indeterminacy in the use of preset criteria for assessment and grading. *Assessment & Evaluation in Higher Education, 34*(2), 159–179. https://doi.org/10.1080/02602930801956059

Richey, R. C., & Klein, J. D. (2014). *Design and development research: Methods, strategies, and issues*. Routledge. Retrieval: https://doi.org/10.4324/9780203826034

Patton, M. Q. (2015). *Qualitative research & evaluation methods: Integrating theory and practice* (4th ed.). Sage Publications. Retrieval: Retrieved from https://uk.sagepub.com/en-gb/eur/qualitative-research-evaluation-methods/book232962

Ruiz-Primo, M. A., & Li, M. (2017). The role of formative assessment in the design and development of measures of teaching quality. In D. M. McInerney & H. K. H. J. Datu (Eds.), *Handbook of formative assessment in a personalised learning context* (pp. 513–536). Springer.

Braun and Clarke (Thematic Analysis): Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp0630a

Ghanbari, N. (2020). Development and validation of a rating scale for Iranian EFL academic writing assessment. *Language Testing in Asia, 10*, 11. https://doi.org/10.1186/s40468-020-00112-3

Peter, S. (2024). Objective assessment criteria reduce the influence of judgment errors based on student characteristics and related subjective associations. *Frontiers in Education, 9*, 1386016. https://doi.org/10.3389/feduc.2024.1386016

Sirimanna, P., Fernando, S., & Jayasinghe, K. (2022). Development of a rating scale for objective assessment in competency-based training programs. *Journal of*

Education and Training Studies, 10(5), 75–88. https://doi.org/10.1002/PMCID-PMC9544375

Depoo, L. (2022). *Factors of quality assessment in higher education and its impact*. ERIC Educational Research & Innovation, 13. https://doi.org/10.13140/RG.2.2.36634.50880

Jones-Devitt, S., & Austen, L. (2021). *A guide to basic evaluation in higher education*. Staffordshire Centre of Learning and Pedagogic Practice. https://www.enhancementthemes.ac.uk/docs/ethemes/about-us/eval-guide_qaas_pdf.pdf

McNamara, D., Ferreira, M., Romkes, M., & Bain, S. (2015). *Rubric development and inter-rater reliability as iterative design* (ERIC No. ED628198). https://files.eric.ed.gov/fulltext/ED628198.pdf

Novak, K. (2014, May 27). *The inter-rater reliability protocol: A must-have for writing assessment.* Novak Education Blog. https://www.novakeducation.com/blog/interrater-reliability-protocol-must-writing-assessment

OECD. (2020). *Education policy evaluation – surveying the OECD landscape.* OECD Publishing. https://doi.org/10.1787/EDU/WKP(2020)24-en

Teaching Commons. (n.d.). *Inter-rater reliability and rubrics.* DePaul University. https://resources.depaul.edu/teaching-commons/teaching-guides/feedback-grading/rubrics/Pages/evaluating-rubrics.aspx