

Development and Validation of an Innovative Performance-Based Assessment Tool

Dianna Rose B. Piñera, RN, LPT¹, Maria Shane S. Del Rosario, RN, PhD²

¹Master of Arts in Nursing, Graduate School, University of Santo Tomas-Legazpi, Legazpi City, Philippines

²Adviser, Graduate School, University of Santo Tomas-Legazpi, Legazpi City, Philippines

Publication history: Received: January 12, 2026; Revised: March 28, 2026; Accepted: April 01, 2026

Email of Corresponding Author: diannarosebelen@gmail.com

Abstract

Critical thinking is a core competency in nursing education because it supports clinical judgment, decision-making, and safe patient care. However, traditional assessment methods in nursing often emphasize procedural accuracy and knowledge recall rather than higher-order reasoning and applied clinical judgment. This study evaluated the existing assessment tool used in NCM 101 – Health Assessment in Nursing and developed and validated an innovative performance-based assessment tool for Level I nursing students at Amando Cope College. A developmental–descriptive research design integrating quantitative and qualitative elements was employed. Data were obtained from 168 Level I nursing students, seven Clinical Instructors, and seven expert validators. Student and instructor feedback were gathered through structured questionnaires using a 4-point Likert scale, while the newly developed instrument was subjected to expert content validation using the Item-Level Content Validity Index (I-CVI) and Scale-Level Content Validity Index (S-CVI). Findings showed that the existing tool was generally perceived as clear, fair, and aligned with course objectives, but it demonstrated limitations in promoting analytical reasoning, higher-order thinking, and differentiation of student performance. In response, a performance-based assessment tool organized by body systems and aligned with the nursing process was developed. Content validation showed strong expert agreement, with I-CVI values ranging from 0.86 to 1.00 and an S-CVI value exceeding the recommended threshold of 0.90. The validated tool offers a more comprehensive, developmentally appropriate, and evidence-based approach to evaluating cognitive, psychomotor, and affective learning outcomes in nursing education.

Keywords: performance-based assessment, nursing education, critical thinking, clinical reasoning, content validity, health assessment

1. Introduction

Critical thinking is widely recognized as an essential competency in nursing because nurses are expected to make sound clinical judgments in increasingly complex and unpredictable healthcare environments. In nursing education, however, traditional evaluation methods often focus on memorization, procedural execution, and theoretical recall rather than the practical application of knowledge in authentic clinical contexts. Recent literature and policy directions in Philippine nursing education have therefore emphasized the need for assessment methods that can capture clinical reasoning, decision-making, and evidence-based problem solving in observable performance.

At Amando Cope College, Level I nursing students enrolled in NCM 101 – Health Assessment in Nursing were primarily assessed through written examinations and return demonstrations. Although these methods supported structured performance of clinical procedures, they were limited in measuring how students interpreted findings, prioritized patient problems, and justified clinical decisions. This created a gap between intended competency-based learning outcomes and the existing approach to student assessment, particularly for novice learners who needed both structure and opportunities to develop higher-order thinking.

In response, the present study was undertaken to evaluate the adequacy of the existing assessment tool, identify its deficiencies in measuring critical thinking, and develop a validated performance-based assessment tool aligned with curricular competencies, the nursing process, and novice-level learning needs. The study aimed to contribute a more authentic, structured, and context-responsive approach to evaluating student performance in health assessment.

2. Material and methods

2.1 Research Design

The study used a descriptive–developmental research design. The descriptive phase examined the adequacy of the existing NCM 101 assessment tool, while the developmental phase focused on the design, refinement, and validation of a new performance-based assessment instrument. This approach allowed the study to move systematically from diagnosis of current practice to evidence-based tool development.

2.2 Locale of the Study

The study was conducted at Amando Cope College in Tabaco City, Albay, Philippines, specifically within the NCM 101 – Health Assessment in Nursing course for Level I nursing students. The institutional setting was selected because the existing assessment practice relied largely on traditional written examinations and return demonstrations, making it a relevant context for tool evaluation and innovation.

2.3 Respondents of the Study

Three groups participated in the study. The first group consisted of 168 Level I nursing students enrolled in NCM 101 during Academic Year 2024–2025. The second group included seven Clinical Instructors who taught or supervised the course. The third group involved seven nursing education experts who served as validators of the newly developed instrument.

2.4 Sample and Sampling Procedure

Total enumeration was used for the student participants, meaning all 168 enrolled Level I nursing students in the course were included. The seven Clinical Instructors were included because of their direct involvement in course assessment. The seven expert validators were chosen through purposive sampling based on their qualifications and experience in nursing education, curriculum development, clinical instruction, and academic leadership.

2.5 Research Instrument

The study used three major materials: the existing NCM 101 assessment tool, structured evaluation forms for expert validators, and the newly developed performance-based assessment tool. The developed instrument was organized by body systems and incorporated the nursing process—assessment, diagnosis, planning, implementation, and evaluation. It was designed to measure observable indicators of critical thinking, including interpretation of findings, prioritization, and justification of clinical decisions. The expert validation form used a 4-point relevance scale, where ratings of 3 or 4 indicated acceptable relevance for CVI computation.

2.6 Data Gathering Procedure

The data gathering process began with collection and review of the existing assessment tools used in NCM 101. Student feedback and Clinical Instructor feedback were then gathered to determine the strengths and limitations of the current assessment approach. Based on these results, a new performance-based assessment tool was developed and submitted to seven expert validators for content validation. The instrument underwent revision based on expert feedback until acceptable content validity and practical applicability were achieved.

2.7 Data Analysis Techniques

Descriptive statistics were used to summarize the responses of students and Clinical Instructors. For the validation phase, content validity measures were employed, specifically the Item-Level Content Validity Index (I-CVI) and the Scale-Level Content Validity Index (S-CVI). Items rated as relevant or highly relevant by experts were retained, while items that fell below the acceptable threshold were revised to improve clarity, specificity, and developmental appropriateness.

2.8 Ethical Considerations

Ethical principles were strictly observed throughout the study. Permission was obtained from the administration of Amando Cope College before implementation. Participation of students, instructors, and validators was voluntary, and all participants were informed of the purpose and procedures of the study. Confidentiality was maintained, no identifying information was included in the presentation of results, and responses did not affect the academic or professional standing of participants.

3. Results and Discussion

3.1 Evaluation of the Current Assessment Tool

Student feedback showed that the current NCM 101 Health Assessment Manual was generally perceived as effective, with an overall mean of 3.29 (Agree). The highest-rated indicator was the feedback mechanism ($M = 3.46$), followed by the tool's ability to challenge critical thinking and help students understand why procedures are important ($M = 3.38$ each). Students also agreed that the tasks reflected realistic clinical situations and that the tool was fair and accurate. These findings indicate that the existing instrument provided structure and instructional value, particularly in guiding foundational skill performance.

Despite these strengths, none of the indicators reached the “Strongly Agree” range. The lowest student ratings were found in clarity of expectations during assessment ($M = 2.98$), adequacy of time allocation ($M = 3.07$), and clarity of evaluation criteria ($M = 3.23$). These results suggest that although the tool was generally acceptable, students still experienced uncertainty about what was expected of them and whether the available time and criteria were sufficient for demonstrating both skills and critical thinking.

3.2 Clinical Instructor Feedback and Identified Deficiencies

Clinical Instructors also rated the existing tool positively overall, with an overall mean of 3.24 (Agree). The strongest areas were alignment with course objectives ($M = 3.86$), clarity of instructions ($M = 3.57$), realism of clinical situations ($M = 3.57$), and consistency of evaluation ($M = 3.57$). These ratings confirm that the current tool was organized and functional for foundational clinical evaluation.

However, Clinical Instructors identified important weaknesses in the tool's ability to assess higher-order reasoning. The lowest ratings were found in promotion of critical thinking beyond procedural steps ($M = 2.71$), support for higher-order thinking skills ($M = 2.71$), and differentiation of student performance based on critical thinking ($M = 2.71$). Additional concern was noted regarding redundancy of tasks ($M = 2.86$) and limited support for analytical and interpretative skills ($M = 3.43$). Taken together with student feedback, these results reveal that the existing manual was strong in procedural guidance but limited in measuring deeper clinical reasoning and judgment.

The identified deficiencies therefore centered on three major issues: inadequate clarity of expectations and criteria, limited support for higher-order thinking, and insufficient differentiation of student performance. These findings are consistent with nursing education literature indicating that traditional checklist-based assessment often captures task completion more effectively than interpretation, prioritization, and decision-making. In this context, refinement of assessment practice was necessary to align evaluation with intended competency-based outcomes.

3.3 Development and Validation of the Performance-Based Assessment Tool

In response to the identified deficiencies, the study developed an innovative performance-based assessment tool organized by body systems and aligned with the nursing process. The tool was designed to incorporate observable reasoning behaviors, clear criteria, and structured performance indicators that would allow instructors to assess not only psychomotor execution but also interpretation, prioritization, planning, and evaluation. This represented a shift from a primarily procedural checklist toward a more comprehensive clinical reasoning instrument.

Content validation findings showed that the developed instrument had strong to excellent content validity. Across the fourteen sections of the tool, most indicators achieved I-CVI values ranging from 0.86 to 1.00, while the S-CVI exceeded the recommended threshold of 0.90. Foundational areas such as General Status and Vital Signs and Pain Assessment showed particularly high expert agreement, while more complex sections such as thorax and lungs, heart and neck vessels, and abdomen showed more variability and therefore required some revision. Items with low agreement were refined rather than removed, resulting in clearer, more measurable, and developmentally appropriate indicators for Level I nursing students.

These validation results suggest that the content of the developed tool was fundamentally sound and that expert disagreement was driven mainly by issues of phrasing, specificity, or level appropriateness rather than conceptual weakness. The integration of the nursing process within each section also strengthened the instrument by linking technical performance with clinical judgment, thereby improving its value as an assessment tool for novice learners.

3.4 Implications for Nursing Education

Comparison of the existing and developed tools showed a clear transformation in assessment design. The previous manual was largely procedure-based, whereas the developed instrument integrated assessment, diagnosis, planning, implementation, and evaluation across body systems. This structural change made the new tool more holistic, clearer in expectations, and more capable of evaluating critical thinking alongside technical performance.

The findings imply that performance-based, behavior-anchored assessment can improve transparency, consistency, and relevance in nursing evaluation. By including explicit indicators for clinical reasoning and decision-making, the validated tool offers educators a more systematic method for assessing learner performance and gives students clearer expectations regarding what constitutes competent practice. The tool is therefore positioned as a stronger alternative to traditional checklist-based assessment in NCM 101 and similar introductory nursing courses.

4. Conclusion & Recommendations

The study concludes that the existing NCM 101 assessment tool was generally effective in guiding structured clinical performance among Level I nursing students, particularly in terms of clarity, fairness, and alignment with course objectives. However, it had important limitations in promoting higher-order thinking, analytical reasoning, and differentiation of student performance. In response, the study successfully developed and validated a performance-based assessment tool that demonstrated strong content validity and stronger alignment with critical thinking, clinical reasoning, and competency-based nursing education.

It is recommended that nursing institutions consider adopting the developed performance-based assessment tool as a validated alternative to traditional checklist-based methods. Clinical Instructors should be oriented and trained in its use, and regular calibration sessions should be conducted to improve inter-rater reliability. Clearer communication of assessment expectations, sample criteria, and better time allocation are also recommended to improve transparency, student confidence, and learner preparedness. Future research may further test the tool across other nursing programs, academic

levels, and assessment platforms.

A limitation of the study is that it was conducted within a single institutional setting and focused on one course among Level I nursing students, which may affect the transferability of findings to other nursing programs. In addition, the study emphasized development and content validation rather than long-term implementation testing, outcome testing, or broader multi-site validation. These limitations suggest the need for further studies that examine the tool's effectiveness over time and across diverse educational settings.

Compliance with ethical standards

This study complied with institutional and academic ethical standards. Permission to conduct the research was secured from Amando Cope College prior to implementation. Participation of students, Clinical Instructors, and expert validators was voluntary, and all respondents were informed of the purpose and procedures of the study. Confidentiality of responses was maintained, data were used solely for academic and research purposes, and no identifying information was included in the reporting of results. Participants were assured that their academic or professional standing would not be affected by participation or non-participation. The author assumes responsibility for the integrity of the manuscript, acknowledges the scholarly sources used in the study, and declares that the work was prepared in accordance with accepted academic and ethical standards.

REFERENCES

- Almanasreh, E., Moles, R., & Chen, T. F. (2020). Evaluation of methods used for estimating content validity. *Research in Social and Administrative Pharmacy*, 16(2), 214–221. <https://doi.org/10.1016/j.sapharm.2019.03.066>
- Boso, C. M., van der Merwe, A. S., & Gross, J. (2021). Students' experiences of assessment in nursing education. *Nurse Education Today*, 102, 104940. <https://doi.org/10.1016/j.nedt.2021.104940>
- Cajanding, R. J. (2021). Clinical reasoning in nursing education: A review of strategies and outcomes. *Belitung Nursing Journal*, 7(6), 456–462. <https://doi.org/10.33546/bnj.1589>
- Cant, R. P., & Cooper, S. J. (2020). Simulation-based learning in nurse education: A systematic review. *Journal of Nursing Education*, 59(5), 281–287. <https://doi.org/10.3928/01484834-20200422-06>
- Carbogim, F. C., Barbosa, A. C., & de Oliveira, L. B. (2021). Educational strategies to promote critical thinking in nursing students: A systematic review. *Nurse Education Today*, 97, 104708. <https://doi.org/10.1016/j.nedt.2020.104708>
- Commission on Higher Education (CHED). (2017). CHED Memorandum Order No. 15, series of 2017: Policies, standards, and guidelines for the Bachelor of Science in Nursing (BSN) program.
- Gaspar, A. D. (2024). Undergraduate student nurses' satisfaction, self-confidence, and critical thinking using simulation-based learning. *Nurse Education Today*, 104, 104981.
- Kim, J., & Park, J. H. (2022). Effects of case-based learning on critical thinking in nursing education. *Nurse Education Today*, 108, 105198. <https://doi.org/10.1016/j.nedt.2021.105198>
- Labrague, L. J., McEnroe-Petitte, D. M., Tsaras, K., Cruz, J. P., Colet, P. C., & Gloe, D. S. (2021). Predictors of clinical competence among nursing students: A cross-sectional study. *Nurse Education in Practice*, 52, 103013. <https://doi.org/10.1016/j.nepr.2021.103013>
- Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research*, 35(6), 382–385. <https://doi.org/10.1097/00006199-198611000-00017>
- Madayag, R. A., et al. (2024). Refining clinical judgment competence in nursing education: A mixed-methods study. *Belitung Nursing Journal*, 10(6), 680–694. <https://doi.org/10.33546/bnj.3560>
- Masuwai, A., Zulkifli, H., & Ahmad, N. (2024). Evaluation of content and face validity of an assessment instrument. *Cogent Education*, 11(1), 2308410. <https://doi.org/10.1080/2331186X.2024.2308410>
- Papathanasiou, I. V., Kleisaris, C. F., Fradelos, E. C., Kakou, K., & Kourkouta, L. (2020). Critical thinking: The development of an essential skill for nursing students. *Acta Informatica Medica*, 28(2), 128–132. <https://doi.org/10.5455/aim.2020.28.128-132>
- Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? *Research in Nursing & Health*, 29(5), 489–497. <https://doi.org/10.1002/nur.20147>
- Polit, D. F., Beck, C. T., & Owen, S. V. (2007). Is the CVI an acceptable indicator of content validity? *Research in Nursing & Health*, 30(4), 459–467. <https://doi.org/10.1002/nur.20199>
- Republic Act No. 9173. (2002). Philippine Nursing Act of 2002.
- Rush, S., Acton, C., Tolley, K., Marks-Maran, D., & Burke, L. (2021). Using simulation in nursing education. *Nurse Education in Practice*, 54, 103136. <https://doi.org/10.1016/j.nepr.2021.103136>
- Shin, S., Park, J. H., & Kim, J. H. (2022). Effectiveness of patient simulation in nursing education. *Nurse Education Today*, 99, 104857. <https://doi.org/10.1016/j.nedt.2020.104857>

- Sullivan, N., et al. (2021). Emerging evidence toward a 2:1 clinical to simulation ratio. *Nurse Education Today*, 96, 104718. <https://doi.org/10.1016/j.nedt.2020.104718>
- Tawfik, A. A., Reeves, T. D., & Stich, A. E. (2020). Guidance in problem-centered learning environments. *Educational Psychology Review*, 32(4), 1123–1156. <https://doi.org/10.1007/s10648-020-09507-6>
- Wang, F. (2024). Content validity and content validity index in instrument development. *Sains Humanika*, 16(1), 1–10.