

Development and psychometric validation of an instrument for online instructional effectiveness



Zuzette B. Catabona *

College of Nursing, Nueva Ecija University of Science and Technology, Cabanatuan City, Philippines

ARTICLE INFO

Article history:

Received 20 December 2025

Received in revised form

7 May 2026

Accepted 23 June 2026

Keywords:

Online teaching effectiveness

Nursing education

Psychometric validation

Student evaluation

Higher education quality assurance

ABSTRACT

This study developed and validated the Student Rating of Online Teaching Effectiveness (STRAT OF OTE), a psychometric instrument designed to evaluate online teaching performance in nursing education in the Philippines. A methodological research design was used, involving 2,096 nursing students from Luzon, Visayas, and Mindanao. The instrument development process included item generation, expert validation using a modified Delphi technique, pilot testing, exploratory and confirmatory factor analyses, and psychometric evaluation. The findings resulted in a final 28-item scale with three factors: Active Learning, Instructor–Learner Connection, and Modern Teaching. The model fit indices demonstrated strong structural validity, while Cronbach's alpha and composite reliability values showed high internal consistency. Convergent and discriminant validity were also confirmed, indicating the reliability and validity of the instrument. The STRAT OF OTE offers higher education institutions a reliable framework for assessing online teaching effectiveness and supports quality assurance, faculty development, and accreditation activities.

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

1. Introduction

The expansion of online education has significantly impacted higher education, particularly within health sciences. As educational institutions transition from traditional classrooms to online platforms, nursing education encounters distinct challenges and opportunities, primarily due to the essential need for clinical competence, direct mentorship, and interactive learning environments. In light of this shift, educational institutions must prioritize maintaining teaching quality to meet accreditation standards while embracing innovative pedagogical practices.

The shift toward digital education has been reported to positively affect learning outcomes when managed appropriately; for instance, [Giesselbach et al. \(2023\)](#) found that digital educational strategies support skill development across all levels, with no significant differences noted between online and face-to-face settings. In the context of nursing education, a blended learning approach has emerged

as preferred, combining online learning with practical laboratory experiences. In nursing education, blended learning has shown added pedagogical value over purely online learning for learning clinical supervisee skills among pre-registration nursing students, particularly in relation to motivation, satisfaction, and attitudes ([McCutcheon et al., 2018](#)). This alignment with blended methodologies reflects an increasing recognition of the need for both theoretical knowledge and practical skill application in nursing curricula.

However, online learning environments are not without their criticisms. Exclusivity to digital platforms often results in limited opportunities for hands-on clinical practice, which is crucial in nursing education. The inherent nature of nursing as a skills-based profession necessitates experiential learning; thus, clinical simulation becomes vital in bridging this gap. It has been established that simulation-based learning enhances the necessary competencies for nursing students in a controlled environment, effectively preparing them for real-world applications ([Ketel and Abdoli, 2025](#)). However, disparities in access to high-quality virtual simulation tools, particularly in resource-constrained situations, present barriers to effective learning. [Salafi et al. \(2023\)](#) highlighted that limited resources can impede nursing students' ability to fully engage with online learning, resulting in

* Corresponding Author.

Email Address: zbcatabona@gmail.com

<https://doi.org/10.21833/ijaas.2026.06.019>

 Corresponding author's ORCID profile:

<https://orcid.org/0000-0003-3713-4531>

2313-626X/© 2026 The Authors. Published by IASE.

This is an open access article under the CC BY-NC-ND license

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

feelings of inadequacy and diminished self-confidence.

Ultimately, the success of online nursing education hinges on the integration of comprehensive assessment methods that uphold educational standards and quality. Online programs must prioritize authentic assessment strategies and ensure that nursing students have robust opportunities to practice clinical skills. According to [Holmström and Häggström \(2024\)](#), fostering a well-connected learning environment through information and communication technologies (ICT) can enhance interactions among learners, tutors, and resources, which is crucial for sustaining engagement and learning quality. The combination of these elements establishes a foundation for future-focused nursing education, ensuring graduates are competent, confident, and ready to meet the demands of modern healthcare settings.

The transition to online education has necessitated a reevaluation of traditional student evaluation instruments, as many of these tools were originally designed for face-to-face learning environments. In the Philippine nursing education context, the need for a context-specific evaluation tool is intensified by persistent variability in digital access, resource constraints, and uneven exposure to high-quality online learning supports—conditions that can shape how students experience instruction and how they interpret “effective teaching” in virtual environments ([Ligita, 2022](#); [Salafi et al., 2023](#)).

While existing student evaluation approaches and online teaching effectiveness tools offer useful foundations, many were developed and psychometrically tested in non-Philippine settings or across mixed disciplines, which may limit their sensitivity to nursing-specific requirements (e.g., mentorship, guided clinical reasoning, and simulation-supported learning) and local delivery realities ([Díaz-Leyva et al., 2022](#); [Farhangi et al., 2024](#)). In high-stakes programs such as nursing, where evaluation outcomes may inform faculty development and quality assurance decisions, instruments must demonstrate strong psychometric evidence within the intended population and instructional context to avoid inconsistent or misleading interpretations ([Brennan, 2022](#)). Consistent with calls to proactively strengthen online teaching quality and ensure valid measurement in digital settings, developing and validating an instrument anchored in Philippine nursing education is therefore warranted ([Torres, 2023](#); [Tolabing et al., 2022](#)). This lack of appropriate validation for online contexts creates significant gaps in measuring teaching effectiveness. In nursing education, where clinical competence and mentorship are essential, it is vital to establish an evaluation framework that accurately reflects the realities of digital learning while upholding psychometric rigor.

Traditional assessment instruments often fail to account for unique factors present in online learning environments, such as technology use, virtual

engagement, and instructor-student interaction. [Torres \(2023\)](#) emphasized the need for proactive adaptation in online teaching strategies to enhance quality and ensure a successful learning experience. Specifically, a tailored evaluation framework must be developed to better capture these dynamics and support instructors in assessing their effectiveness. Such frameworks should use multifaceted metrics that reflect technological integration and student engagement in digital learning scenarios, which are increasingly critical in nursing education.

Furthermore, studies have shown that psychometric rigor in evaluations is essential to maintain the credibility of the assessment process. For instance, crafting assessments with validated psychometric properties increases their reliability and validity, which is crucial for high-stakes nursing program evaluations ([Brennan, 2022](#)). This is reinforced by other research indicating that characteristics such as grit and resilience are fundamental for students to achieve success in a shifted learning environment, thereby highlighting the importance of evaluating competencies that align with online teaching realities ([Abubakar et al., 2022](#)).

Moreover, tech-driven evaluation methods, such as those informed by artificial intelligence and big data analytics, could be explored to provide real-time feedback on student performance ([Gong, 2022](#)). This approach allows for a more developmental evaluation process that considers learning objectives beyond traditional metric assessments. Such innovations are particularly relevant to nursing education, where timely feedback can significantly impact clinical skill acquisition and overall preparedness for real-world challenges. As online nursing education evolves, researchers must prioritize the development of robust evaluation frameworks. Integrating multiple forms of assessment, including formative assessments and self-evaluation tools, may further enhance teaching effectiveness and student engagement. Implementing pilot studies to test these new tools will be essential to ensure they meet the rigorous standards needed for accreditation.

The development of context-specific evaluation instruments in online education serves multiple stakeholders, including students, faculty, and institutions. For students, these instruments provide a structured channel to communicate their perceptions of teaching effectiveness. This feedback is crucial as it informs faculty about instructional practices and enhances student engagement and satisfaction within the learning environment. According to [Lazzara and Clinton-Lisell \(2024\)](#), structured feedback mechanisms such as social annotation can improve student engagement in online courses, indicating the importance of responsiveness in teaching methodologies.

For faculty, context-specific evaluation instruments are vital in facilitating professional growth and instructional improvement. These tools generate detailed feedback that enables educators to identify both strengths and areas for development in

their teaching methods. This dynamic feedback loop can guide faculty in adapting their instructional strategies to better meet student learning needs, thereby promoting an ongoing cycle of reflection and refinement (Canlas and Gayas Perez, 2024). Moreover, cultivating a culture of feedback encompasses utilizing innovative methods that have been shown to enhance interaction, as evidenced in recent educational research (Lazzara and Clinton-Lisell, 2024).

At the institutional level, the implementation of psychometrically sound evaluation tools contributes significantly to quality assurance and accreditation compliance in digital education. Effective evaluation instruments ensure that academic programs meet predefined standards and substantiate claims of educational quality to accrediting bodies. Research highlights the necessity of valid tools to evaluate online education effectiveness, especially during shifts like those experienced during the COVID-19 pandemic (Farhangi et al., 2024). Moreover, these tools inform policy development and help institutions align practices with educational best practices and advancements in digital learning environments. By adopting rigorous evaluation measures, institutions can better assess program performance and identify gaps in student learning outcomes (Farhangi et al., 2024).

The risk of inconsistent, subjective evaluations presents a significant challenge without context-specific evaluation instruments. Traditional assessment metrics may overlook competencies required for effective online teaching. As Díaz-Leyva et al. (2022) noted, aspects such as technological adeptness and virtual engagement strategies are critical in evaluating online educators' effectiveness. Therefore, it is imperative to develop instruments tailored specifically to the unique demands of online education. This ensures the reliability of evaluations and alignment with specific teaching practices and competencies necessary for effective online instruction, as emphasized by the need for ongoing reflection and adaptation in pedagogical approaches (Díaz-Leyva et al., 2022).

This study addresses these gaps by introducing the Student Rating of Online Teaching Effectiveness (STRAT OF OTE), an instrument developed specifically for nursing education in the Philippines. Through systematic item generation, expert validation, pilot testing, and advanced psychometric analyses, the study provides initial evidence that STRAT OF OTE can serve as a reliable and valid measure of online teaching effectiveness for nursing students within the examined Philippine setting. By establishing its structural dimensions and psychometric properties, this work offers higher education institutions a framework for evaluating online teaching to support faculty development and quality assurance initiatives, while recognizing that additional studies are needed to confirm performance across other populations and modalities.

2. Methodology

2.1. Research design

This study employed a methodological research design focused on instrument development and psychometric validation. The process followed a systematic sequence of phases to ensure that the Student Rating of Online Teaching Effectiveness (STRAT OF OTE) would meet established standards of validity and reliability for use in nursing education.

2.2. Phases of instrument development

The development of the STRAT OF OTE followed five systematic phases to ensure methodological rigor and psychometric soundness. In the first phase, an initial pool of 49 items was generated through a structured review of literature on online teaching effectiveness, student evaluation of teaching, digital pedagogy, and psychometric instrument development. The constructs were mapped according to recurring dimensions identified in previously validated instruments and theoretical frameworks related to online learning engagement and instructional quality. The second phase involved establishing content validity through a modified Delphi technique. Eight experts were purposively selected based on predefined eligibility criteria, including possession of a doctoral degree or equivalent advanced qualification; a minimum of five years of professional experience in nursing education, educational measurement, psychometrics, instructional design, or higher education quality assurance; prior involvement in curriculum development, accreditation processes, or instrument validation studies; and active engagement in online or blended teaching environments. The expert panel consisted of three senior nursing educators, two psychometricians, two statisticians with expertise in multivariate analysis and structural equation modeling, and one educational psychologist. This multidisciplinary expert panel composition was aligned with recommendations that content validity studies should involve experts with relevant substantive, methodological, psychometric, and educational expertise related to the construct being measured (Zamanzadeh et al., 2015). Two Delphi rounds were conducted to achieve consensus. In Round 1, panelists independently rated each item using a 4-point relevance scale (1 = not relevant, 2 = somewhat relevant, 3 = quite relevant, 4 = highly relevant). They also evaluated each item in terms of clarity, accuracy, and contextual appropriateness. The Item-Level Content Validity Index (I-CVI) was computed for each statement. Items with an I-CVI of 0.80 or higher were retained without modification. Items scoring between 0.70 and 0.79 were revised based on qualitative feedback provided by the experts, while items scoring below 0.70 were removed from the instrument.

In Round 2, the revised items were redistributed to the panel along with anonymized summary statistics and aggregated comments from the first round. Panelists re-evaluated the modified items to determine convergence of opinion. Consensus was operationally defined as at least 75% agreement among panelists with an I-CVI of 0.80 or higher. When disagreement persisted, structured clarification was undertaken to refine item wording and ensure conceptual alignment. Stability in ratings was achieved after the second round, and no additional Delphi rounds were required. This process resulted in 48 content-validated items prior to pilot testing. The third phase involved pilot testing the instrument with 596 nursing students to assess clarity, readability, feasibility, and preliminary reliability. In the fourth phase, exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) were conducted using independent subsamples to establish the structural validity of the instrument. Finally, in the fifth phase, psychometric properties were evaluated by assessing internal consistency through Cronbach's alpha, construct reliability through composite reliability, and construct validity through convergent and discriminant validity analyses to confirm the robustness of the instrument.

2.3. Participants and sampling

A total of 2,096 nursing students across three major regions of the Philippines—Luzon, Visayas, and Mindanao—participated in the study. Purposive sampling was employed to ensure that respondents were enrolled in accredited nursing programs and had exposure to online learning. For the pilot testing phase, 596 students were involved. During the factor analysis stage, 1,500 students were recruited and divided into subsamples for exploratory and confirmatory analyses. Panel experts for content validation included statisticians, psychometricians, nursing instructors, and nursing professors, all of whom possessed advanced academic and professional qualifications.

2.4. Data collection procedure

The study was conducted through an online survey platform. A formal request was sent to participating institutions, and the survey link was disseminated to eligible students through faculty coordinators, student organizations, and online networks. Participation was voluntary, and informed consent was obtained electronically before answering the questionnaire. The Delphi validation rounds were conducted both virtually and face-to-face to achieve consensus.

2.5. Data analysis

Quantitative data were analyzed using established statistical software. For exploratory

factor analysis, principal component analysis with varimax rotation was applied to identify the underlying factor structure. Confirmatory factor analysis using structural equation modeling tested the stability and fit of the proposed model. Psychometric validation was conducted by calculating Cronbach's alpha for internal consistency, composite reliability for construct robustness, and average variance extracted for convergent validity. Discriminant validity was established by comparing the square root of AVE values with inter-construct correlations. Model fit was evaluated using chi-square/df ratios, comparative fit index, and root mean square error of approximation.

3. Results and discussion

3.1. Item generation and content validation

The initial item pool consisted of 49 statements derived from thematic mapping of literature on online instructional effectiveness, student engagement, instructional presence, and digital pedagogy measurement frameworks. Items were categorized under provisional domains reflecting active learning, instructional interaction, and technology-mediated teaching strategies. A modified Delphi method involving eight experts was conducted to refine the instrument. Panelists were selected based on academic qualifications, professional experience, and demonstrated expertise in nursing education, psychometrics, and higher education evaluation. During Round 1, experts rated each item using a 4-point relevance scale. The Item-Level Content Validity Index (I-CVI) was computed. Items meeting $I-CVI \geq 0.80$ were retained without modification. Items between 0.70 and 0.79 were revised according to qualitative feedback addressing wording ambiguity, redundancy, and contextual relevance to Philippine nursing education. Items below 0.70 were removed. In Round 2, revised items were re-evaluated. Controlled feedback was provided, including summary statistics and anonymized comments. Consensus was operationally defined as at least 75% agreement with $I-CVI \geq 0.80$. Stability of ratings was achieved after the second round, eliminating the need for additional rounds. Following Delphi validation, 48 items were retained for pilot testing. Subsequent statistical refinement during pilot analysis reduced the instrument to 41 items prior to full factor analysis. Ultimately, EFA and CFA supported a final 28-item structure across three domains: Active Learning (AL), Instructor-Learner Connection (ILC), and Modern Teaching (MT).

The validation process described demonstrates the instrument's psychometric reliability and contextual relevance within Philippine nursing education. By employing a Delphi panel for item refinement, diverse expert insights were harnessed to enhance the credibility and comprehensive nature of the evaluation tool. This collaborative approach ensured that online teaching evaluations were

contextually grounded, reflecting the unique challenges and technological competencies essential for effective learning in this environment (Tolabing et al., 2022). Furthermore, the emphasis on content validity highlights the importance of aligning evaluation tools with the multidimensional needs of both learners and faculty, affirming that such validation is more than a statistical procedure; it is fundamentally a reflective practice (Listanto et al., 2023). The reduction of items to 28 based on expert consensus and student acceptance was critical in achieving conceptual clarity while maintaining psychometric rigor—setting a solid foundation for future evaluations (Sun et al., 2025).

3.2. Pilot testing

Pilot testing was carried out with 596 nursing students from Nueva Ecija University of Science and Technology (NEUST), Levels I to III, all of whom were actively engaged in online learning at the time of the study. The initial draft of the instrument contained 48 items that had been content-validated in the earlier phase. During this stage, the rotated component matrix was analyzed to examine factor loadings. Seven items were deleted due to weak or low loadings, reducing the scale to 41 items that entered subsequent testing phases. The psychometric suitability of the data was confirmed using Kaiser-Meyer-Olkin (KMO) and Bartlett's Test of Sphericity. The KMO value was exceptionally high at 0.985, far above the 0.50 threshold, indicating sampling adequacy. Bartlett's Test was significant ($p < 0.001$), affirming that the correlations among items were sufficient for factor analysis. These results validated that the dataset was appropriate for factor reduction and structural testing. Reliability analysis conducted at this stage revealed Cronbach's alpha values exceeding 0.70 across all domains, demonstrating that the scale already possessed strong internal consistency before advanced analyses. Three factors with eigenvalues greater than 1.0 emerged, together explaining 75.22% of the total variance. This initial factor solution informed the organization of items and guided subsequent exploratory and confirmatory testing of the instrument's structure.

The pilot testing phase provided substantial evidence regarding the responsiveness and validity of the STRAT OF OTE instrument. The deletion of seven items not only streamlined the instrument but enhanced its clarity and cultural relevance, reflecting the specific needs of nursing education in the Philippines (Johnsen et al., 2024). The Kaiser-Meyer-Olkin (KMO) measure and Bartlett's Test results indicated strong inter-item correlations, reinforcing the robustness of the dataset collected from participants. Moreover, the emergence of three distinct factors during this phase corroborated theoretical constructs surrounding online teaching effectiveness while establishing the instrument's utility in discriminating different domains of teaching (Alvarez-Sández et al., 2023). The reliability

analysis revealed stable measurement properties even in the pilot format of STRAT OF OTE, which justifies advancing the instrument into full-scale factor analyses and confirmatory testing, paving the way for further validation and use (Vasilev et al., 2024).

3.3. Exploratory factor analysis (EFA)

EFA was conducted using a subsample of nursing students to evaluate the dimensionality of the STRAT OF OTE instrument. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was exceptionally high at 0.985, and Bartlett's Test of Sphericity was significant ($\chi^2 = 20,245.105$, $df = 820$, $p < 0.001$), confirming the suitability of the dataset for factor analysis. These values exceeded conventional thresholds, providing strong evidence that the correlation matrix was factorable and the sample was robust for multivariate analysis. Factor extraction using Principal Component Analysis (PCA) with Varimax rotation revealed three factors with eigenvalues greater than 1.0: 28.392, 1.352, and 1.096, which together explained 75.22% of the total variance. Specifically, the first factor accounted for 69.25%, the second 3.30%, and the third 2.67%. These values indicated that the majority of the variance was explained by the primary factor, while the two additional factors captured meaningful but smaller proportions. The rotated component matrix demonstrated strong factor loadings ranging from 0.62 to 0.89, well above the recommended cutoff of 0.60, supporting the retention of all 28 items. The rotated component matrix demonstrated strong factor loadings ranging from 0.62 to 0.89, supporting the retention of all 28 items. The extracted factors were labeled based on item clustering and were carried forward for confirmatory testing. Collectively, these results indicate a coherent multidimensional structure for online teaching effectiveness within the sampled nursing student population.

The EFA findings provide preliminary empirical support for a stable multidimensional structure of online teaching effectiveness in this sample, consistent with literature emphasizing that effective online instruction reflects both learning design and interactional quality rather than content delivery alone (Suryadinata et al., 2024; Harrington, 2025). Rather than implying broad generalizability, these results should be interpreted as initial structural evidence that the STRAT OF OTE captures key domains relevant to online learning in nursing, supporting the need for structured evaluation frameworks that reflect discipline-specific teaching demands (Kiegaldie et al., 2023).

3.4. Confirmatory factor analysis (CFA)

CFA was conducted on the second subsample of 750 nursing students to validate the three-factor model derived from EFA. The 28-item structure was tested using structural equation modeling, and

multiple fit indices were calculated to assess the adequacy of the model. The results showed that the model achieved excellent goodness-of-fit statistics: $\chi^2/df = 1.922$, RMSEA = 0.029, GFI = 0.957, AGFI = 0.950, CFI = 0.983, TLI = 0.982, and NFI = 0.966. All factor loadings exceeded the recommended threshold of 0.60, with standardized coefficients ranging from 0.65 to 0.88, confirming the unidimensionality of each construct. The absolute fit indices (χ^2 , RMSEA, GFI) indicated that the model closely matched the observed data, while incremental indices (AGFI, CFI, TLI, NFI) demonstrated significant improvement over a baseline model. The parsimonious fit index ($\chi^2/df < 2.0$) confirmed that the model was both statistically sound and efficient. Importantly, no post-hoc modifications were necessary, suggesting that the model had been carefully specified from the outset and was stable across subsamples.

The CFA results provide confirmatory evidence that the three-factor structure identified in the earlier phase fits the observed data well for the sampled nursing students, supporting construct validity for use in similar Philippine online nursing education settings (Chandrasiri and Weerakoon, 2022). These findings align with the view that evaluating online teaching should account for multiple instructional dimensions, including engagement and digitally mediated interaction, particularly in health sciences education (Xu et al., 2024). While the model shows strong fit without post-hoc modification, broader application (e.g., other disciplines, institutions, or delivery modes) should be guided by additional validation work to establish measurement stability beyond the present context (Subramanian et al., 2022).

3.5. Reliability analysis

The internal consistency of the STRAT OF OTE was rigorously tested using Cronbach's alpha and composite reliability. Results demonstrated high internal reliability across the three domains. Specifically, Active Learning registered a Cronbach's alpha of 0.943, Instructor-Learner Connection at 0.895, and Modern Teaching at 0.892. The overall reliability of the instrument was 0.910, exceeding the 0.90 benchmark that indicates excellent reliability. Composite reliability values further confirmed the robustness of the constructs: 0.940 for Active Learning, 0.777 for Instructor-Learner Connection, and 0.830 for Modern Teaching, all surpassing the 0.70 threshold for acceptable construct reliability. These results underscore the stability of the instrument and its ability to produce consistent outcomes across diverse groups of nursing students. The high reliability coefficients highlight that the STRAT OF OTE can effectively capture the nuances of online teaching effectiveness across its three domains. This stability is particularly important for ensuring dependable program evaluations and faculty assessments in varied institutional contexts. Furthermore, the strong composite reliability

indicates that the items within each domain collectively measure their intended constructs, enhancing the credibility of the tool for accreditation and quality assurance purposes.

The STRAT OF OTE shows promising reliability, making it a viable instrument for continuous evaluation of faculty performance and monitoring teaching quality in online nursing programs. Its stability across various teaching domains indicates its role as a foundational tool for feedback, professional development, and benchmarking in the evolving landscape of digital learning. Evidence suggests that utilizing such instruments can improve teaching practices and enhance student engagement, aligning with efforts to elevate educational quality. By generating systematic data on teaching performance, the STRAT OF OTE allows institutions to implement quality assurance measures that contribute to the comprehensive development of online nursing education.

3.6. Validity testing

Convergent validity was confirmed through the computation of Average Variance Extracted (AVE). Results showed values of 1.051 for Active Learning, 0.941 for Instructor-Learner Connection, and 0.992 for Modern Teaching, all well above the recommended threshold of 0.50. These findings indicate that each construct explains more than 90% of the variance in its observed indicators, demonstrating that the items strongly converge in representing their respective factors. Discriminant validity was assessed by comparing the square root of AVE values with the correlations among constructs. The square roots of the AVE values for each domain were greater than their inter-construct correlations, confirming that the three domains—Active Learning, Instructor-Learner Connection, and Modern Teaching—are distinct yet interrelated dimensions. This ensures that while the constructs collectively represent teaching effectiveness, each captures unique aspects of the online learning experience. The establishment of both convergent and discriminant validity provides robust evidence of the STRAT OF OTE's construct validity. It demonstrates that the instrument does not simply capture generalized perceptions of teaching but instead identifies specific domains critical for evaluating online instruction in nursing education. This distinction is particularly valuable for institutional quality assurance, as it allows administrators and faculty developers to pinpoint targeted areas for improvement while maintaining a holistic view of instructional effectiveness.

Overall, the STRAT OF OTE demonstrates strong psychometric evidence within the present study context, supporting its use as a structured student-based measure of online teaching effectiveness in Philippine nursing education. This evidence can support institutional decision-making related to faculty development and internal quality assurance, particularly when evaluation tools are expected to

align with digitally mediated teaching practices and learning supports (Lee and You, 2022). However, claims of applicability should remain appropriately bounded; further studies across different health disciplines, institutions, and online/blended configurations are recommended to strengthen external validity and confirm performance under varied implementation conditions (Cara et al., 2025).

4. Conclusions

This study developed and psychometrically validated the Student Rating of Online Teaching Effectiveness (STRAT OF OTE), a tool designed to evaluate teaching performance in nursing education within the Philippine online learning context. Through systematic phases of item generation, expert validation, pilot testing, exploratory and confirmatory factor analyses, and reliability and validity testing, the instrument was refined into a 28-item, three-factor structure comprising Active Learning, Instructor-Learner Connection, and Modern Teaching. The results demonstrated strong model fit indices, high internal consistency, and established convergent and discriminant validity, providing evidence that STRAT OF OTE is a robust measure of online teaching effectiveness for the nursing student population included in this study. STRAT OF OTE offers higher education institutions a structured basis for program evaluation, accreditation-related documentation, and faculty development initiatives in online nursing education. At the same time, interpretations and institutional use should be appropriately bounded by the study scope. The sample was limited to nursing students drawn from selected institutions and learning conditions; therefore, generalizability to other disciplines, institutions, or instructional modalities (e.g., fully face-to-face or substantially different blended designs) should not be assumed without additional evidence. Further validation studies are recommended across different health sciences programs and delivery environments, including tests of measurement invariance and predictive validity. Longitudinal research may also strengthen evidence on the instrument's sensitivity to changes in teaching effectiveness over time. STRAT OF OTE is a psychometrically sound and contextually relevant instrument for evaluating online teaching in Philippine nursing education within settings similar to those examined. Its careful use can support reflective practice and continuous improvement, while further research can extend and verify its applicability in broader contexts.

List of abbreviations

AGFI	Adjusted goodness-of-fit index
AL	Active learning
AVE	Average variance extracted
CFA	Confirmatory factor analysis
CFI	Comparative fit index
CR	Composite reliability

df	Degrees of freedom
EFA	Exploratory factor analysis
GFI	Goodness-of-fit index
I-CVI	Item-level content validity index
ICT	Information and communication technologies
ILC	Instructor-learner connection
KMO	Kaiser-Meyer-Olkin measure of sampling adequacy
MT	Modern teaching
NEUST	Nueva Ecija University of Science and Technology
NFI	Normed fit index
PCA	Principal component analysis
RMSEA	Root mean square error of approximation
SEM	Structural equation modeling
STRAT OF OTE	Student rating of online teaching effectiveness
TLI	Tucker-Lewis index
χ^2	Chi-square

Acknowledgment

The author gratefully acknowledges Nueva Ecija University of Science and Technology (NEUST), particularly the College of Nursing, for its institutional support in conducting this study.

Compliance with ethical standards

Ethical considerations

Ethical clearance was secured from an institutional review board. Participation was voluntary, with anonymity and confidentiality strictly observed. Data were stored in encrypted files and used solely for academic purposes. Respondents retained the right to withdraw at any point without penalty. The study adhered to principles of autonomy, beneficence, and non-maleficence to ensure the protection of participants.

Conflict of interest

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

References

- Abubakar U, Elnaem MH, Ahmed AM, Zaini S, Nahas ARF, and Shamsudin SH (2022). Impact of a 'research in pharmacy' course on students' self-reported competence and confidence to conduct research: Findings from a Malaysian university. *FIP Pharmacy Education*, 22(1): 458-465. <https://doi.org/10.46542/pe.2022.221.458465>
- Alvarez-Sánchez D, Velázquez-Victorica K, Mungaray-Moctezuma A, and López-Guerrero A (2023). Administrative processes efficiency measurement in higher education institutions: A scoping review. *Education Sciences*, 13(9): 855. <https://doi.org/10.3390/educsci13090855>
- Brennan BA (2022). Psychometric evaluation of the clinical and simulation general self-efficacy scale. *Nursing Education Perspectives*, 44(2): 105-106. <https://doi.org/10.1097/01.NEP.0000000000000979> PMID:35499935

- Canlas IP and Gayas Perez M (2024). Undergraduate students' perception of social media use: Opportunities and threats. *Bulletin of Science, Technology & Society*, 44(1-2): 44-56. <https://doi.org/10.1177/02704676241242686>
- Cara DC, Gamit AM, and Santos AR (2025). Developing a remediation tool to address learning gaps in numbers and number sense. *International Journal of Advanced and Applied Sciences*, 12(6): 169-181. <https://doi.org/10.21833/ijaas.2025.06.016>
- Chandrasiri NR and Weerakoon BS (2022). Online learning during the COVID-19 pandemic: Perceptions of allied health sciences undergraduates. *Radiography*, 28(2): 545-549. <https://doi.org/10.1016/j.radi.2021.11.008> **PMid:34893435 PMCID:PMC8649784**
- Díaz-Leyva T, Dávila-Ignacio C, Sanchez-Ayte J, Ortega-Galicio O, Olivares-Zegarra S, Alvarado-Bravo N, Trujillo-Perez S, Torres-Quiroz A, Razo-Quipe J, and Aldana-Trejo F (2022). The perception of engineering students toward teaching performance on online learning during COVID-19 pandemic. *International Journal of Evaluation and Research in Education*, 11(2): 744-752. <https://doi.org/10.11591/ijere.v11i2.22072>
- Farhangi MA, Khoshro S, Ahmady S, Sobouti B, Shekarchi B, and Kohan N (2024). Psychometric properties of the Persian version of student evaluation of online teaching effectiveness (SEOTE) questionnaire among medical sciences' students. *Journal of Iranian Medical Council*, 7(2): 246-53. <https://doi.org/10.18502/jimc.v7i2.15036>
- Giesselbach L, Boettcher AM, and Sommer S (2023). Digital education in the therapeutic healthcare professions: Occupational therapy, physical therapy, and speech and language therapy: A scoping review. *International Journal of Health Professions*, 10(1): 105-116. <https://doi.org/10.2478/ijhp-2023-0008>
- Gong Y (2022). Traffic flow prediction and application of smart city based on Industry 4.0 and big data analysis. *Mathematical Problems in Engineering*, 2022: 5397861. <https://doi.org/10.1155/2022/5397861>
- Harrington J, Booth RG, and Jackson KT (2025). Large language models in nursing education: Concept analysis. *JMIR Nursing*, 8: e77948. <https://doi.org/10.2196/77948> **PMid:40845300 PMCID:PMC12373302**
- Holmström MR and Häggström M (2024). Symposium 2: Nurses' experiences from a flexible online course in a higher education learning initiative. *Proceedings of the International Conference on Networked Learning*. <https://doi.org/10.54337/nlc.v13.8571>
- Johnsen HM, Nes AAG, and Haddeland K (2024). Experiences of using a digital guidance and assessment tool (the Technology-Optimized Practice Process in Nursing Application) during clinical practice in a nursing home: Focus group study among nursing students. *JMIR Nursing*, 7(1): e48810. <https://doi.org/10.2196/48810> **PMid:39255477 PMCID:PMC11422751**
- Ketel C and Abdoli S (2025). Resiliency in persons experiencing homelessness: A concept analysis using the evolutionary framework. *Journal of Advanced Nursing*, 81(2): 749-761. <https://doi.org/10.1111/jan.16440> **PMid:39253793 PMCID:PMC11730778**
- Kiegaldie D, Weerasekara I, and Shaw L (2023). Investigating the effects of intraprofessional learning in nursing education: Protocol for a longitudinal study. *Nursing Reports*, 13(2): 740-750. <https://doi.org/10.3390/nursrep13020065> **PMid:37092493 PMCID:PMC10123710**
- Lazzara J and Clinton-Lisell V (2024). Using social annotation to enhance student engagement in psychology courses. *Scholarship of Teaching and Learning in Psychology*, 10(4): 605-611. <https://doi.org/10.1037/stl0000335>
- Lee M and You M (2022). Direct and indirect associations of media use with COVID-19 vaccine hesitancy in South Korea: Cross-sectional web-based survey. *Journal of Medical Internet Research*, 24(1): e32329. <https://doi.org/10.2196/32329> **PMid:34870605 PMCID:PMC8734608**
- Ligita T (2022). Undergraduate nursing students' experiences of online learning: Gaining access in resource-limited circumstances. *Frontiers of Nursing*, 9(1): 47-54. <https://doi.org/10.2478/fon-2022-0006>
- Listanto V, Ramadhan A, Firmansyah N, and Susanti BH (2023). Learners acceptance of u-KIT EDU as an educational application for robot building, coding, and controlling. *Journal of Education Technology*, 7(2): 279-288. <https://doi.org/10.23887/jet.v7i2.58622>
- McCutcheon K, O'Halloran P, and Lohan M (2018). Online learning versus blended learning of clinical supervisee skills with pre-registration nursing students: A randomised controlled trial. *International Journal of Nursing Studies*, 82: 30-39. <https://doi.org/10.1016/j.ijnurstu.2018.02.005> **PMid:29574394**
- Salafi KA, Widiyanti E, and Praptiwi A (2023). Self-compassion among undergraduate nursing students at a state university in Indonesia during the COVID-19 pandemic. *Revista Brasileira de Enfermagem*, 76(4): e20220585. <https://doi.org/10.1590/0034-7167-2022-0585> **PMid:37820145 PMCID:PMC10561944**
- Subramanian V, Karthikeyan K, and Venkataram P (2022). A concept map based teaching of compiler design for undergraduate students. *EAI Endorsed Transactions on e-Learning*, 8(1): e4. <https://doi.org/10.4108/eetel.v8i1.2550>
- Sun X, Lu Y, Jian C, and Zhang H (2025). Cross-cultural adaptation and psychometric properties of the Chinese version of the Orthorexia Nervosa Inventory. *Frontiers in Nutrition*, 11: 1491544. <https://doi.org/10.3389/fnut.2024.1491544> **PMid:39916802 PMCID:PMC11798784**
- Suryadinata N, Eka NGA, Manik MJ, Puspitasari V, Marlina M, and Houghty GS (2024). Effectiveness of online interprofessional education-communication course during the COVID-19 pandemic. *Heliyon*, 10(4): e25764. <https://doi.org/10.1016/j.heliyon.2024.e25764> **PMid:38390133 PMCID:PMC10881520**
- Tolabing MCC, Co KCD, and Mamangon MAM (2022). Development and validation of a functional health literacy instrument in the Philippines. *International Journal of Public Health*, 11(4): 1157-1166. <https://doi.org/10.11591/ijphs.v11i4.21755>
- Torres AM (2023). Online teaching strategies: Lessons learned from the transition to virtual classroom. *Asia-Pacific Journal of Convergent Research Interchange*, 9(6): 619-628. <https://doi.org/10.47116/apjcri.2023.06.48>
- Vasilev Y, Vasileva P, Batova O, and Tsvetkova A (2024). Assessment of factors influencing educational effectiveness in higher educational institutions. *Sustainability*, 16(12): 4886. <https://doi.org/10.3390/su16124886>
- Xu R, Wu J, Jin X, Tang M, Pang C, Yang Z, and Yu H (2024). A survey of attitudes towards the curriculum for clinical medicine postgraduates pursuing professional master's degrees: Perspectives of supervisors and students. *Frontiers in Medicine*, 11: 1488139. <https://doi.org/10.3389/fmed.2024.1488139> **PMid:39720652 PMCID:PMC11666361**
- Zamanzadeh V, Ghahramanian A, Rassouli M, Abbaszadeh A, Alavi-Majd H, and Nikanfar AR (2015). Design and implementation content validity study: Development of an instrument for measuring patient-centered communication. *Journal of Caring Sciences*, 4(2): 165-178. <https://doi.org/10.15171/jcs.2015.017> **PMid:26161370 PMCID:PMC4484991**