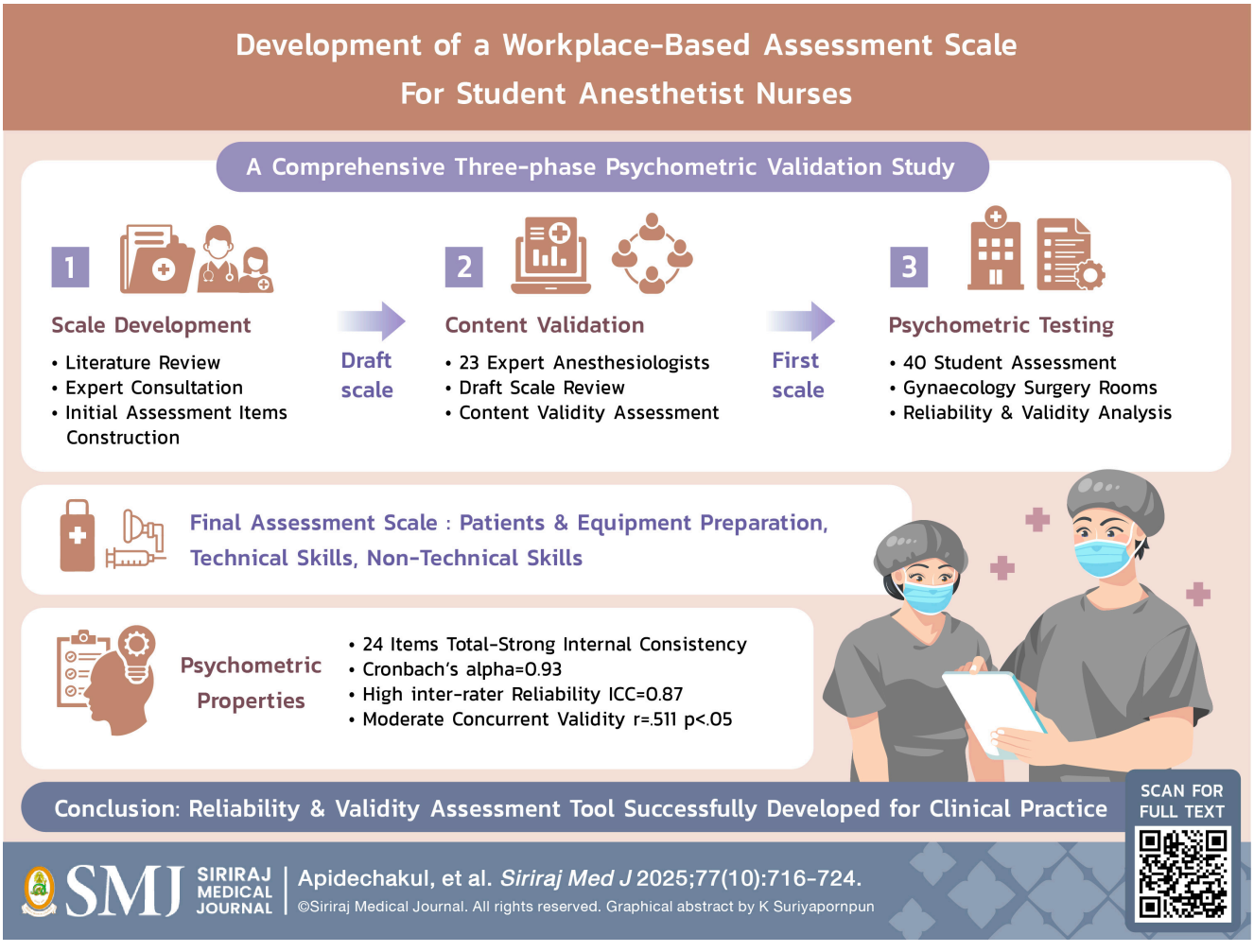
Development of a Workplace-based Assessment Scale for Student Anesthetist Nurses

Parichad Apidechakul, Ph.D.[1], Namtip Triyasunant, M.D.[2], Ladda Permpolprasert, B.Sc.[2], Wiruntri Punchuklang, B.Sc.[2], Phongthara Vichitvejpaisal, M.D., Ph.D.[2,*]

[1]*Siriraj Health Science Education Excellence Center, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand,* [2]*Department of Anesthesiology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand.*



*Corresponding author: Phongthara Vichitvejpaisal
E-mail: phongthara@gmail.com
Received 24 June 2025 Revised 8 August 2025 Accepted 8 August 2025
ORCID ID:http://orcid.org/0000-0003-2656-6718
https://doi.org/10.33192/smj.v77i10.276186

ABSTRACT

Objective: To develop and validate a skills assessment tool for evaluating the workplace performance of student anesthetist nurses through a psychometric analysis of its reliability and validity.

Materials and Methods: The study was conducted in three sequential phases: Scale Development: A comprehensive literature review and expert consultations were undertaken to construct the initial assessment items. Content Validation: Three expert anesthetists reviewed the draft scale for content validity. Psychometric Testing: The final scale was evaluated using data from 40 student anesthetist nurse assessments in 2023-2024. Recruitment was conducted in the gynecology surgery rooms.

Results: The finalized workplace-based assessment scale consisted of 24 items spanning three key domains: patient and equipment preparation, technical skills, and non-technical skills. The psychometric analysis demonstrated strong internal consistency (Cronbach's $\alpha = 0.93$), high inter-rater reliability (The Intraclass Correlation Coefficient (ICC) = 0.87), and moderate concurrent validity, as evidenced by Pearson's correlation analysis ($r=.511$, $p<.05$, $N=40$).

Conclusion: The newly developed workplace-based assessment scale exhibits strong psychometric properties, providing a reliable and comprehensive tool for evaluating the clinical competencies of student anesthetist nurses in practice settings.

Keywords: Medical education; scale; student anesthetist nurses; workplace-based assessment (Siriraj Med J 2025; 77: 716-724)

INTRODUCTION

Healthcare professionals prioritize patient safety in their practice. The Ministry of Public Health and the Institute for Healthcare Quality Accreditation jointly considered the criteria of the Institute for Healthcare Improvement (IHI) and the Joint Commission on Accreditation of Healthcare Organizations (JCAHO) to select safety issues that are consistent with the Thai context and to put these ideas into practice for drafting policies (Patient and Personnel Safety, 2P Safety Goals), with the primary objective of fostering the potential and safety consciousness of healthcare personnel. The policy covers anesthesiology duties, ensuring anesthesia is provided safely. Training sessions focus on risk management, guided by the Royal College of Anesthesiologists of Thailand (RCAT), emphasizing non-technical skills through lectures and hands-on training for student anesthetist nurses.[1]

Professional nurses in the anesthesia nurse program receive comprehensive training focused on patient safety. They learn cognitive processes, patient monitoring, and procedural skills under the supervision of anesthesiologists during training. Over the one-year course, knowledge, skills, and attitudes are taught and evaluated in theory and practice. Additionally, the curriculum consists of practical exams with real-life patients in operating theaters and simulated crisis scenarios using high-fidelity manikins. In actuality, intubating patients under general anesthesia is the responsibility of most student anesthetist nurses during training. Furthermore, when they graduate as anesthetist nurses, general anesthesia is the primary task nurses must complete. Thailand still has a dearth of anesthesiologists and an increased general anesthesia workload, with nearly 1 million cases yearly handled by anesthetist nurses.[2] Student anesthetist nurses in our training institute located in Thailand stated their opinions after taking a practical exam in emergency general anesthesia situations with obstetric simulated patients. More than 70 percent did not meet the passing criteria. This may happen due to real-life training sessions within the operating room; nursing students do not have the opportunity to make their own decisions. Routine work must be performed closely under the supervision of the anesthesia staff. As a result, nursing students may not recognize their decision-making abilities. In training, nursing students also have limited opportunities to perform general anesthesia on cesarean section patients due to patient safety concerns. Regarding the primary obligation of nurse anesthetists in performing intubation, the process requires vigilance in maintaining airway patency, preventing aspiration, and preventing gastric contents from entering the lungs.

This anesthesia technique requires skilled anesthesia staff to evaluate patients before the procedure, especially in emergency surgical situations, to prevent complications and unintended events. Appropriate intubation and anesthesia administration skills enable a quick recovery after the procedure, allowing for endotracheal tube removal.[3] Therefore, anesthetist nurses play a crucial

role in patient care and assist doctors during the pre-anesthesia, intra-anesthesia, and post-anesthesia periods.

Anesthesiology includes teams of anesthesiologists crucial for delivering anesthesia services, ensuring patient safety and comfort during procedures. The Nursing and Midwifery Council's competency framework for anesthetist nurses has been collated and summarized into a list of anesthesia nurse competencies,[4] which include: 1) Preoperative planning; 2) Intraoperative management and patient monitoring; 3) Anesthesia nursing practice during the recovery period; 4) Decision-making during crises; 5) Clinical pharmacology of drugs used in anesthesia; 6) Airway management in an unconscious patient; 7) Knowledge management and using technology and information in evidence-based practice, and 8) Communicating with the anesthesiologist or surgeons. Practical tools to assess student anesthetist nurse competencies in general anesthesia are rare.

Clinical skills are considered the most critical aspect of patient care.[5] The assessment of clinical competence is challenging and complex. There is a trend toward a competency-based training process that emphasizes formative assessment over summative assessment. Workplace-based Assessment (WPBA) scale is an essential part of the evaluation that involves direct observation of clinical and non-clinical skills with actual patients in a particular workplace to provide constructive feedback for the trainees.[6] Mini-clinical examination exercise (mini-CEX) is a form of on-the-job assessment where a specialist or instructor observes the trainee performing on-site work with patients. Students are graded on history taking, physical exam skills, counseling, and feedback. The mini-CEX has been developed for use in anesthesia training in many countries.[7-9] Applying these assessment methods in accordance with the nurse anesthetist competency framework to student anesthetist nurses sounds interesting, especially in real practice settings with actual patients.

Non-technical skills have been utilized in industry and other agencies for some time. A majority of anesthesia incidents occur during work. Investigations of adverse events have shown that up to 80% of cases are caused by human error factors such as ineffective communication, inadequate patient supervision, and a failure to inspect medications and equipment. This is not just a lack of technical knowledge. The Anesthesiologist Non-Technical Skills (ANTS) is a system developed by industrial psychologists and anesthesiologists, covering topics like task management, teamwork, situational awareness, and decision-making. Combining non-technical skills with medical knowledge and clinical skills supports safe and effective performance in daily work and emergencies. ANTS can be observed in good anesthesia practice.[10] The performance assessment tools used to guide evaluation should be clear and transparent. These tools can assess individual behavior, provide information for the training process, and constructive feedback on skill development. Non-clinical skills can be identified through observable behaviors.

The general anesthesia practical skills assessment form of the anesthesia nurse training program at Siriraj Hospital was originally a rubric assessment with four levels of performance in the workplace. The scores were weighted according to the importance of the workload before, during, and after general anesthesia, which lasted for 6 months. The scores that appear on the form are complicated in that they must be calculated in terms of decimal points for each level of the score. The original form was developed and used for more than 5 years for formal evaluation. Behavioral observation in real-world settings has patient context, location, and complexity variability. The researcher is thus interested in developing a new form to assess the practical skills of nursing students in preoperative, intraoperative, and post-anesthesia evaluation for the ability to perform airway opening and intubation and incorporate elements of the ANTS System in assessment. As such, this new form will provide feedback on both technical and non-technical skills to student anesthetist nurses, resulting in the daily development of anesthesia nurse competencies. This study aims to develop and validate a skills assessment tool for evaluating the workplace performance of student anesthetist nurses through a psychometric analysis of its reliability and validity.

MATERIALS AND METHODS

Setting and design

This is a descriptive cross-sectional study conducted from July to October 2024. The study population were student anesthetist nurses, training in the Faculty of Medicine Siriraj Hospital, Mahidol University, Thailand.

Ethical approval

The study received ethical approval from the Human Research Ethics Committee of Siriraj Institutional Review Board (SIRB) at the Faculty of Medicine Siriraj Hospital Certificate of Approval No. Si 522/2024. It was based on the Declaration of Helsinki,[11] the Belmont Report, the International Conference on Harmonization in Good Clinical Practice (ICH-GCP), and the International Guidelines for Human Research, along with Thai laws and regulations.

Sampling criteria

The sample size for a bivariate correlation was determined using power analysis. The power analysis was conducted in a software program as G*Power[12], using an alpha of 0.05, a power of 0.95, and a medium effect size (0.05); the required sample size was 38. The Faculty of Medicine at Siriraj Hospital trains 40 anesthetist nurses each year. This demonstrates the program's dedication to maintaining a focused cohort size because of the limited number of students accepted annually. Therefore, the researcher determined that the final sample size for this research would be 40 individuals in the academic year 2023-2024. Recruitment was conducted in the gynecology surgery rooms. The inclusion criteria of nurse students and patients were: 1) Being a student anesthetist nurse at an accredited training center by the Royal College of Anesthesia of Thailand; 2) Having completed theoretical and practical training in general anesthesia for patients over 18; and for patients 3) Gynecologic patients aged > 18 years with American Society of Anesthesiologists (ASA) physical status 1-2. The exclusion criteria were: 1) Student anesthetist nurses without experience or training in general anesthesia; and for patients 2) Patients with ASA classification ≥ 3 or those younger than 18.

Data collection

Anesthesia training program chairpersons were contacted to share the study's background and gauge their willingness to participate. Posters were distributed via student groups for recruitment assistance. After participants were recruited for additional assessments beyond standard evaluations, explicit, informed consent was provided for their participation. All eligible participants had consented beforehand, understanding they could withdraw at any time. They were assured of data confidentiality and anonymity. Valid clinical and non-clinical competency assessments in authentic practice settings are considered crucial for student anesthetist nurses. The research was comprised of three steps, as shown in Fig 1.

Step 1: Scale development	
• Item generation	• Literature review • Previous research
24 draft items	
Step 2: Content validation	
Expert opinion: A panel of two anesthesiologists and one expert in measurement and evaluation in education.	• Item-Objective Congruence (IOC) calculated - 1st round - 2nd round
Drafted scale (24 items)	
• Field test 40 student anesthetist nurses	• Student anesthetist nurses performed general anesthesia to patients undergoing gynecological surgery with a risk assessment score of no more than level 2. • Two trained raters observed the same student for a duration of 1 to 4 hours
Final draft scale (24 items)	
Step 2: Psychometric testing	
• Criterion-related validity	• Pearson's correlation test (n=40) • Correlated with existing 6-month clinical rotation competency scores before, during, and after administering evaluations that utilized standard practical examination methods
• Reliability	• Internal consistency (n=40) Cronbach's alpha for the scale is >0.70
• Inter-rater reliability	• Intraclass correlation coefficients (ICC) >0.80
The workplace-based assessment scale for student anesthetist nurses has 24 items and 3 subscales.	

Fig 1. A workplace-based assessment scale for student anesthetist nurses

Step 1: Scale development

The goal was to create practical assessment tools to evaluate the student anesthetist nurses' clinical and non-technical skills, foster their development, and enhance the quality of their training. The technical and non-technical skills items were developed based on the competency framework for anesthetist nurses as determined by the Nursing and Midwifery Council of Thailand and the Anesthetists' Non-Technical Skills (ANTS). Subsequently, psychometric tests were performed to establish the reliability and validity of the scale.

Step 2: Content validation

Specific performance indicators were carefully formulated, and their congruence with conceptual definitions was evaluated by two anesthesiologists and one expert in measurement and evaluation in education using the Item-Objective Congruence index (IOC). IOC of < 0.50 were modified.[13] This rigorous process established the content validity of the newly developed workplace assessment tools. Then, a total of 24 items were prepared to assess anesthetist nurse competencies in the workplace with a sample group.

Step 3: Psychometric testing

Psychometric testing involves implementing and validating these workplace-based assessments. Preliminary items were scored on a four-point scale, ranging from 1 (fair) to 4 (excellent). The newly developed instruments underwent field testing to assess the competencies of student anesthetist nurses who provided general anesthesia to patients undergoing gynecological surgery with a risk assessment score of no more than level 2 according to the American Society of Anesthesiologist guidelines under the supervision of the attending anesthesiologist throughout the general anesthesia period, following the standard procedures for general anesthesia of the Department of Anesthesiology, Faculty of Medicine Siriraj Hospital, in the gynecology surgery room, (approximately 2-4 hours). Two trained raters observed the same student before, during, and after the administration of general anesthesia. Assessment results from the workplace-based assessment tools were analyzed against existing 6-month clinical rotation competency scores before, during, and after administering evaluations that utilized standard practical examination methods. This comprehensive two-phase study provided an evidence-based approach to developing valid and reliable workplace-based assessment tools for anesthesia nurse competencies.

Statistical analysis

IBM SPSS statistics, version 29 (IBM Corp, Armonk, NY, USA) was used to analyze the collected data. The student demographic data were presented using descriptive statistics. The criterion-related validity of the new workplace assessments was analyzed by comparing their results to those from the established examination process at a statistical significance level of 0.05, using Pearson's correlation coefficient to check validity. Validity coefficient values above 0.35 were interpreted as very beneficial.[14] Additionally, the internal consistency reliability of the workplace assessment instruments was calculated using the Cronbach's alpha coefficient, and a Cronbach's α of > 0.70 was indicated to be acceptable.[15] The intra-class correlation coefficient was calculated as inter-rater reliability. The student competencies were assessed by two trained anesthetist nurse raters simultaneously. Intra-class correlation coefficients ranging from 0.75 to 1.00 show consistent inter-rater scoring from good to very good.[15]

RESULTS

The present study gathered data on students' age, sex, and work experience with general anesthesia.

General characteristics of the sample

A total of 40 nurse anesthetist students participated in the study. Most participants were women (95%); the mean age was 28.7 ± 1.9 years. The average number of cases experienced in general anesthesia was 142.7 ± 10.2 cases, as shown in Table 1.

TABLE 1. Demographic data and details regarding the participants (N=40).

Demographic characteristics	Value
Sex: Female	38 (95%)
Age	28.7 ± 1.9
Number of cases experiences in general anesthesia	142.7 ± 10.2

Data presented as a number (percentage) or mean ± SD

Step 1: Scale development

A comprehensive literature review and expert consultations were undertaken to construct the initial assessment items. The new WPBA scale had 24 items spanning three key domains: patient and equipment preparation (10 items), technical skills (10 items), and non-technical skills (4 items). Preliminary items were scored on a four-point scale, ranging from 1 (fair) to 4 (excellent).

Rating Scale:

Excellent: Demonstrates outstanding behavior and fully meets expectations.

Good: Shows appropriate behavior with minor areas for improvement.

Fair: Displays acceptable behavior but with occasional inconsistencies.

Poor: Exhibits behavior that is below the expected standard and considered unacceptable.

Responses marked as "N/A" (Not Available) required evaluation for validity, reliability, and alignment with the research objectives. In this study, listwise deletion was applied to remove entire cases containing "N/A" responses when the rate was below 5–10% and the missingness appeared to be random. This approach helped maintain data quality and prevent score inflation.[16]

The Nursing and Midwifery Council's competency framework for anesthetist nurses has been collated and summarized into a list of anesthesia nurse competencies4 and elements of the ANTS System in assessment.[9] (Supplementary file 1)

Step 2: Content validation

Three experts reviewed the content validity of the scale, and the scale achieved an index of Item-Objective Congruence (IOC) ranging between 0.67 and 1.0. Two items did not fulfill the criteria due to unclear content, so the researcher revised them (IOC<0.5).

Step 3: Psychometric testing

The final scale was evaluated using data from 40 student anesthetist nurse assessments. To further check the quality of the new scale, we used Pearson's correlation coefficient, and the criterion-related validity of the workplace scale was investigated against those from the established examination process with a prior scale at a statistical significance level of 0.05. The correlation coefficient for the new scale compared to the old formal scale was 0.511. ($r=0.511$, $p<0.05$, $n=40$) As shown by Pearson's correlation analysis, in Fig 2, moderate concurrent validity was indicated.

Furthermore, the internal consistency reliability of the workplace assessment instruments was calculated using the Cronbach's alpha coefficient. The psychometric analysis demonstrated strong internal consistency (Cronbach's $\alpha = 0.93$).

The intra-class correlation coefficient was calculated as inter-rater reliability. Two trained anesthetist nurse raters assessed student competencies simultaneously—intra-class correlation coefficients ranging from 0.75 to 1.00 show consistent inter-rater scoring from good to very good.[17] This study, the reliability of scoring consistency between the two raters, as the Intraclass Correlation Coefficient (ICC), was calculated for their evaluations at the 95% confidence level. The type of ICC used includes the model (2-way mixed effects), type (mean of multiple raters), and the definition (consistency or agreement). An ICC (3, 2) where "3" signifies a two-way mixed-effects

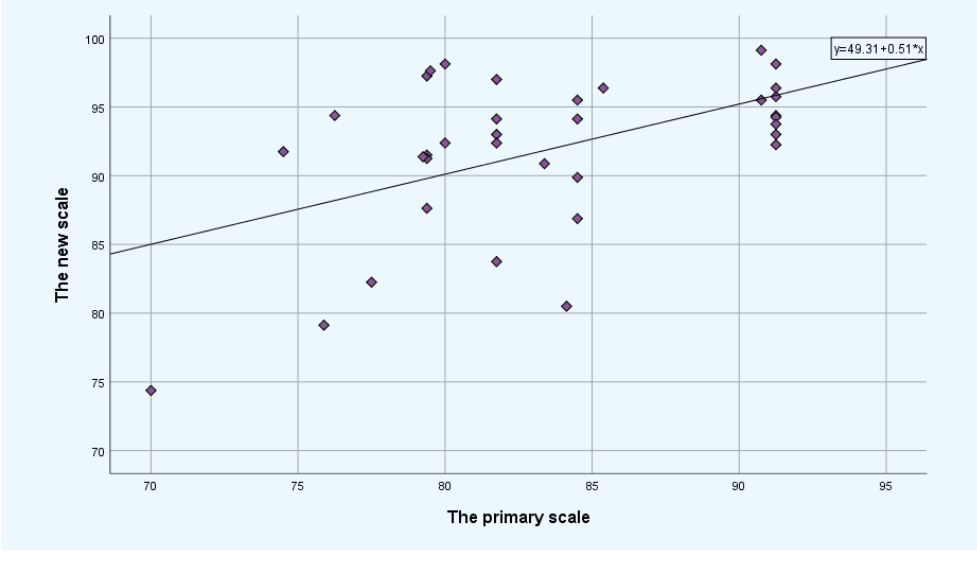


Fig 2. Correlation for the new scale and the primary scale scores

model where raters are considered fixed, and reliability is assessed based on the mean of multiple two raters or measurements.[18] The new assessment scale has an average two-way mixed effects model ICC (3,2) value of 0.93 (0.87 - 0.97), indicating high dependability as in Table 2.

DISCUSSION

The finalized WPBA scale consisted of 24 items spanning three key domains: patient and equipment preparation and technical and non-technical skills. The psychometric analysis demonstrated strong internal consistency (Cronbach's α = 0.93), high inter-rater reliability (ICC = 0.93), and moderate concurrent validity, as evidenced by Pearson's correlation analysis. These findings affirm the scale's robustness in reliably assessing clinical performance.

The development and validation of WPBA scales for student anesthetist nurses addresses a significant gap in competency assessment within anesthesia education. This study contributes to the existing literature by developing a context-specific assessment tool capturing the unique competencies required in anesthesia nurse practice. Our findings provide insights regarding the psychometric properties of the developed scale, with notable strengths and limitations.

Validity Evidence: This WPBA scale demonstrates strong content validity established through expert review and alignment with professional frameworks. WPBA instruments should be grounded in professional standards while remaining sensitive to specific educational contexts.[19] Examination of the professional judgment and decision-making of strength and conditioning coaches' workplace assessments evaluates competence dimensions beyond cognitive knowledge, including contextual judgment and professional behaviors.[20] This WPBA scale is aimed at assessing the practical skills of nursing students in preoperative, intraoperative, and post-anesthesia phases and incorporates elements of the ANTS System in assessment. The anesthetist experts reviewed the draft scale (IOC<0.5) for content validity, and finally, the Item-Objective Congruence (IOC) was over 0.66. The criterion-related validity was evidenced by moderate to strong correlations with existing performance measures, similar to findings regarding anesthesia technical skills assessment.[21]

The improved framework and updated items of the new scale are likely to capture additional dimensions, which may account for the moderate—rather than high—correlation observed.

The psychometric analysis of the new scale demonstrated strong internal consistency (Cronbach's alpha = 0.93), in line with the results of similar studies; previous studies reported this amount for the whole instrument in the range of 0.916 to 0.975.[22] Similar challenges have been reported in achieving high reliability when assessing rarely encountered clinical scenarios.[23]

The inter-rater reliability of two trained anesthetist nurses showed high dependability. The high inter-rater reliability shows that the WPBA scale yields consistent measurements across various evaluators, which is essential for establishing the scale's psychometric properties. This consistency indicates that the assessment criteria are clear and objective enough for trained evaluators to reach similar conclusions when evaluating the same performance. Delfino AE et al. (2023)[24] documented similar challenges in achieving consistent ratings of non-technical skills. Structured rater training improved ICC values by approximately 0.12 across domains. Therefore, training enhances assessment reliability but cannot eliminate subjective judgment elements.[25] It is recommended that other institutions offering training, including those in Thailand, adopt this approach. Consequently, an appendix should be appended to the evaluation form utilized, enabling all individuals who review it to consider its further application.

TABLE 2. Intraclass Correlation Coefficient (ICC).

Measures	Intraclass Correlation Coefficient: ICC (95% CI)	p
One rater	0.88 (0.77 - 0.94)	p< 0.001
Average value from 2 raters	0.93 (0.87 - 0.97)	p< 0.001

Two-way random effects model where both people effects and measures effects are random

Limitation

A primary limitation is the lack of longitudinal validity evidence demonstrating the scale's predictive relationship with post-graduation performance,[26] documented correlations between simulation-based assessments, and subsequent practice. However, the relationship between WPBA scores and practice outcomes requires further research.

The initially planned effect size of r = 0.05 was too small for this study; a more appropriate effect size would have been in the range of r = 0.30–0.50.[27] Additionally, a post-hoc power analysis revealed that, with n = 40 and an observed effect size of r = 0.511, the study achieved approximately 99% power (α = 0.05, two-tailed)—well above the conventional 0.80 threshold. This significant observed effect suggests a meaningful relationship, despite the study's sample size limitations. Future research should employ larger sample sizes to enhance generalizability and improve the precision of effect size estimates.

The scale emphasizes observable behaviors but may not fully capture the cognitive processes underlying clinical decisions. Incorporating cognitive assessment methods could improve this. The think-aloud protocols could enhance clinical reasoning evaluation in future iterations.[28] Several assessment domains remain inherently subjective despite structured rating scales.[29]

CONCLUSION

The workplace assessment scale for student anesthetist nurses shows strong psychometric properties in validity and reliability. Its strengths include content validity, internal consistency, and educational impact, while challenges exist in inter-rater reliability for non-technical skills and feasibility in complex clinical scenarios. Future refinement should aim at technology integration for improved assessment efficiency, longitudinal validation for predictive validity, and adaptation to enhance generalizability across various contexts.[30] These improvements would boost the effectiveness of workplace assessments in anesthesia nurse education, ultimately promoting better patient care through efficient competency evaluation.

Data Availability Statement

Dataset 1. Raw response data of participants to each item of the new workplace-based assessment scale.

Dataset 2. Raw response data of participants to each item of the primary workplace-based assessment scale.

ACKNOWLEDGEMENTS

The authors thank the Siriraj Health Science Education Excellence Center, Faculty of Medicine Siriraj Hospital, for funding our research.

DECLARATIONS

Grants and Funding Information

This study was funded by Siriraj Health Science Education Excellence Center, Faculty of Medicine Siriraj Hospital, Mahidol University, Thailand, under grant number R016861009.

Conflict of Interest

The authors declare no conflict of interest.

Registration Number of Clinical Trial

None.

Author Contributions

Conceptualization and methodology, P.A., N.T., and P.V.; Investigation, L.P., W.P.; Formal analysis, P.A. and P.V.; Visualization and writing – original draft, P.A.; Writing – review and editing, P.A., N.T., and P.V.; Funding acquisition, P.A., L.P., W.P.; Supervision, P.V. All authors have read and agreed to the final version of the manuscript.

Use of Artificial Intelligence

The authors used Grammarly AI to assist with grammar correction and sentence refinement. The authors thoroughly validated and approved all AI-assisted content to ensure accuracy and compliance with academic and ethical standards.

REFERENCES

1. Antonsen KN, Hofstad JK. Simulating preoperative preparations with focus on non-technical skills in an OR nursing education program in Norway. In: Springer International Publishing; 2022. p. 37–52.
2. Ratanasuwan P, Sriraj W, Punjasawadwong Y, Choorat J, Charuluxananan S, Pravitharangul T. The Perioperative and Anesthetic Adverse Events in Thailand (PAAd Thai) study of anesthetic equipment malfunction or failure: An analysis of 2206 incident reports. J Med Assoc Thai. 2021;104:286–92.
3. Smedley P, Nicholas-Holley J. Recovery of patients from anaesthesia and surgery. In: Rodger D, Henshaw K, Rawling P, Miller S, editors. Fundamentals of Operating Department Practice. Cambridge: Cambridge University Press; 2022. p. 384–406.
4. Nucharoenkul S. Competencies and certification of knowledge and specialized expertise for advanced practice nurses. 1st ed. Bangkok: Siriyod Printing; 2010.
5. Matlhaba K. Essential clinical skills. In: Enhancing Clinical Competence of Graduate Nurses. Cham: Springer Nature Switzerland; 2025. p. 113–34.
6. Boker AM. Toward competency-based curriculum: Application of workplace-based assessment tools in the National Saudi Arabian Anesthesia Training Program. Saudi J Anaesth. 2016;10(4):417–22.
7. Khalafi A, Sharbatdar Y, Khajeali N, Haghighizadeh MH, Vaziri M. Improvement of the clinical skills of nurse anesthesia

students using mini-clinical evaluation exercises in Iran: A randomized controlled study. J Educ Eval Health Prof. 2023;20:12.
8. Mortaz Hejri S, Jalili M, Masoomi R, Shirazi M, Nedjat S, Norcini J. The utility of mini-clinical evaluation exercise in undergraduate and postgraduate medical education: A BEME review: BEME Guide No. 59. Med Teach. 2020;42(2):125–42.
9. Kurdi MS, Hungund BR. Evaluation of mini-clinical evaluation exercise (mini-CEX) as a method for assessing clinical skills in anaesthesia postgraduate education. Indian J Anaesth. 2021; 65(3):248–52.
10. Flin R, Fletcher G, Glavin R, Maran N, Patey R. Anaesthetists' Non-Technical Skills (ANTS) System Handbook. Aberdeen: University of Aberdeen; 2003.
11. World Medical Association. World Medical Association Declaration of Helsinki: Ethical principles for medical research involving human subjects. JAMA. 2013;310:2191–4.
12. Faul F, Erdfelder E, Buchner A, Lang AG. Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. Behav Res Methods. 2009;41(4):1149–60.
13. Kanjanawasee S. Classical test theory. 7th ed. Bangkok: Chulalongkorn University Press; 2013.
14. Saad S, Carter GW, Rothenberg M, Israelson E. Testing and assessment: An employer's guide to good practices. Washington, DC: U.S. Department of Labor; 1999.
15. Taber KS. The use of Cronbach's alpha when developing and reporting research instruments in science education. Res Sci Educ. 2018;48:1273–96.
16. Popovich D. How to treat missing data in survey research. J Mark Theory Pract. 2025;33(1):43-59.
17. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med. 2016;15(2):155–63.
18. Senthil Kumar VS, Shahraz S. Intraclass correlation for reliability assessment: The introduction of a validated program in SAS (ICC6). Health Serv Outcomes Res Methodol. 2024;24(1): 1–3.
19. Chin M, Pack R, Cristancho S. "A whole other competence story": Exploring faculty perspectives on the process of workplace-based assessment of entrustable professional activities. Adv Health Sci Educ Theory Pract. 2023;28(2):369–85.
20. Slabbert B, Hoole C. A competency framework for coaches working in coaching development centres. SA J Ind Psychol. 2021;47(1):1–10.
21. Wingate TG, Bourdage JS, Steel P. Evaluating interview criterion-related validity for distinct constructs: A meta-analysis. Int J Sel Assess. 2025;33(1):1–21.
22. Gümüş MM, Kukul V. Developing a digital competence scale for teachers: Validity and reliability study. Educ Inf Technol. 2023;28(3):2747–65.
23. Maj M, van Os J, De Hert M, Gaebel W, Galderisi S, Green MF, et al. The clinical characterization of the patient with primary psychosis aimed at personalization of management. World J Psychiatry. 2021;20(1):4–33.
24. Delfino AE, Fuentes-López EA, de la Fuente RF, Altermatt FR. Cross-cultural adaptation and validation of the Spanish version of the Anesthetists' Non-Technical Skills (ANTS) assessment tool. J Clin Anesth. 2023;84:1-5.
25. Yeates P, Maluf A, Kinston R, Cope N, Cullen K, Cole A, et al. A realist evaluation of how, why and when objective structured clinical exams (OSCEs) are experienced as an authentic assessment of clinical preparedness. Med Teach. 2025;47(3):458–66.
26. Yi L, Shuai T, Zhou J, Cheng L, Jiménez-Herrera MF, Tian X. Development and validation of a machine learning-based predictive model for compassion fatigue in Chinese nursing interns: A cross-sectional study utilizing latent profile analysis. BMC Med Educ. 2024;24(1):1–20.
27. Martin EL, Martinez DA. The effect size in scientific publication. Educ. XXI.2023;26(1):9-17.
28. Gamborg ML, Mehlsen M, Paltved C, Vetter SS, Musaeus P. Clinical decision-making and adaptive expertise in residency: A think-aloud study. BMC Med Educ. 2023;23(1):1–20.
29. McIntosh TR, Liu T, Susnjak T, Watters P, Halgamuge MN. A reasoning and value alignment test to assess advanced GPT reasoning. ACM Trans Interact Intell Syst. 2024;14(3):1–37.
30. Chudapongse R, Sangpetngam B, Chankaew E, Aurboonyawat T, Churojana A, Withayasuk P. Comparative Evaluation of Imaging Modalities for Eligibility in Endovascular Treatment of Delayed Onset Acute Anterior Circulation Ischemic Stroke in Siriraj Hospital: A Retrospective Analysis. Siriraj Med J. 2024; 76(9):573-80.