

Authentic Writing Assessment In Nursing Education, Low-Tech Digital Assessment In Resource-Constrained Nursing Classrooms

Handayani. SB¹, Yenni Rozimela², Syahrul Ramadhan³

E-mail: handayani1976sofyan@gmail.com,

¹Keperawatan, Akper Kesdam I/BB Padang, Indonesia

²³FBS, Universitas Negeri Padang, Indonesia

Abstrak

Penelitian ini mengembangkan dan mengevaluasi asesmen menulis autentik digital berbasis teknologi sederhana untuk mahasiswa English for Nursing pada lingkungan pembelajaran dengan keterbatasan sumber daya. Berlandaskan teori English for Specific Purposes, asesmen autentik, serta prinsip kegunaan asesmen bahasa, produk ini menekankan aspek validitas, reliabilitas, autentisitas, kepraktisan, dan dampak pendidikan. Penelitian ini menggunakan desain Educational Research and Development dengan model ADDIE yang dimodifikasi, meliputi tahap analisis, desain, pengembangan, validasi ahli, implementasi terbatas, dan evaluasi. Uji coba melibatkan 37 mahasiswa English for Nursing dan tiga validator yang terdiri atas satu ahli bahasa Inggris dan dua ahli pendidikan keperawatan. Data dikumpulkan melalui analisis dokumen, analisis kebutuhan, lembar validasi, tes menulis, angket kepraktisan, dan umpan balik pengguna. Data kuantitatif dianalisis menggunakan Aiken's V, Cronbach's alpha, intraclass correlation coefficients, persentase kepraktisan, normalized gain, paired-samples t-test, dan Cohen's d. Hasil analisis kebutuhan menunjukkan tingkat kebutuhan yang sangat tinggi dengan nilai rata-rata keseluruhan sebesar 4,43. Asesmen yang dikembangkan memperoleh rata-rata Aiken's V sebesar 0,90, Cronbach's alpha sebesar 0,891, ICC single measure sebesar 0,82, dan ICC average measure sebesar 0,93. Tingkat kepraktisan menurut mahasiswa mencapai 88,2%, sedangkan kepraktisan menurut ahli dan pengguna mencapai 91,2%. Skor menulis mahasiswa meningkat dari 68,84 menjadi 79,27, yang menunjukkan efektivitas awal yang menjanjikan, meskipun penelitian ini belum menggunakan kelompok kontrol dan masih terbatas pada uji coba di satu institusi.

Kata Kunci: *Asesmen Autentik, Bahasa Inggris Bagi Keperawatan, Keterampilan Menulis, Digital Berbasis Teknologi Sederhana, Pendidikan Keperawatan*

Abstract

This study developed and evaluated a low-tech digital authentic writing assessment for English for Nursing students in resource-constrained learning environments. Grounded in English for Specific Purposes theory, authentic assessment, and language assessment usefulness principles, the product emphasized validity, reliability, authenticity, practicality, and educational impact. An Educational Research and Development design was employed using a modified ADDIE model covering analysis, design, development, expert validation, limited implementation, and evaluation. The pilot involved 37 English for Nursing students and three validators, consisting of one English language expert and two nursing education experts. Data were collected through document analysis, needs analysis, validation sheets, writing tests, practicality questionnaires, and user feedback. Quantitative data were analyzed using Aiken's V, Cronbach's alpha, intraclass correlation coefficients, practicality percentages, normalized gain, paired-samples t-test, and Cohen's d. The needs analysis showed very high demand, with an overall mean of 4.43. The assessment achieved an average Aiken's V of 0.90, Cronbach's alpha of 0.891, ICC single measure of 0.82, and ICC average measure of 0.93. Student practicality reached 88.2%, while expert-user practicality reached 91.2%. Writing scores improved from 68.84 to 79.27, indicating promising preliminary effectiveness despite the absence of a control group and its single-institution pilot scope within this initial study.

Keywords: *Authentic Assessment, English For Nursing, Writing Skills, Low-Tech Digital, Nursing Education*

Introduction

English is no longer merely an extra subject in nursing education, but a professional skill needed to support clinical, academic, and workplace communication. Nursing students use English to write clinical notes, understand patient data, read medical references, and interact with patients, families, and healthcare professionals. This need becomes more urgent because nursing graduates may work in clinical and global health contexts where English often facilitates professional communication. However, English instruction for nursing students is still frequently too general and does not fully reflect the actual communication tasks found in clinical learning and healthcare settings. Recent studies Fisma, (2024); Marleni et al., (2023); Widanta et al., (2023) argued that English for Nursing should be developed based on English for Specific Purposes principles because students need language learning that reflects academic, clinical, and workplace demands. Among the four English skills, writing deserves particular attention because written communication connects clinical reasoning, professional accountability, and linguistic accuracy. Nursing students must write patient summaries, nursing care plans, referral notes, health education messages, reflective reports, and basic clinical documentation. These tasks grammar control, professional tone, accurate medical terminology, clear organization, and context-sensitive correctness.

Therefore, English writing assessment for nursing students should measure their ability to perform authentic nursing-related writing tasks (Huseynova, 2023; Khojasteh et al., 2025; Orfanò & Wingate, 2024). In this context, authentic assessment becomes highly relevant because it allows writing tasks to represent situations that students may encounter in clinical learning, hospital communication, patient education, and interdisciplinary collaboration. Authentic writing assessment does not ask students to produce isolated sentences or general essays. Instead, it requires them to write texts that resemble real academic and professional tasks in nursing contexts. Vlachopoulos & Makri (2024) argued that authentic assessment in higher education uses real-world tasks to evaluate students' knowledge, skills, and attitudes. Zaim et al., (2020) also showed that authentic assessment requires clear language functions, relevant activities, and explicit scoring rubrics. In line with this view, Sekarsari et al. (2023) emphasized that authentic assessment makes language evaluation more meaningful when classroom tasks reflect communicative performance. Thus, an authentic English writing assessment for nursing students should evaluate not only language quality but also nursing-specific communicative appropriateness. Therefore, an authentic English writing assessment for nursing students should measure not only grammar, vocabulary, coherence, and mechanics, but also clinical relevance, accuracy of patient information, appropriate medical terminology, professional tone, ethical sensitivity, and audience awareness. Despite the growing recognition of English for Nursing, authentic assessment, and digital learning tools, a clear gap remains in the availability of validated writing assessment instruments that directly measure nursing-related English writing performance. Previous studies have discussed nursing students' English needs, ESP curriculum design, learning resources, speaking communication, authentic assessment, and digital writing technologies (Astutik et al., 2024; Fisma, 2024; Orfanò & Wingate, 2024; Widanta et al., 2023).

Orfanò & Wingate (2024) emphasize that nursing students need explicit guidance in disciplinary writing conventions, while (Rohmani et al., 2024) Rohmani et al. (2024) illustrate the value of structured digital support for nursing documentation. Sekarsari et al. (2023) further suggest that authentic assessment improves the relevance of classroom evaluation when criteria reflect communicative performance. Accordingly, the present study conceptualizes nursing writing as an integrated construct involving language, genre, professional context, and task purpose.

Meanwhile, digital assessment has become increasingly important in higher education because it can improve data management, feedback delivery, scoring documentation, and learning access. However, many institutions cannot depend on expensive platforms, advanced automated scoring systems, or high-bandwidth digital environments. This condition is especially relevant in nursing education contexts where budgets, laboratories, internet stability, and lecturer training may vary across institutions. Hu et al., (2025) report that digital technologies have been used in authentic assessment across higher education, but their value depends on purposeful design and implementation. Astutik et

al. (2024)) show that automated writing evaluation can support writing classrooms, yet it also raises issues related to accuracy, teacher roles, and feedback interpretation. Therefore, a low-tech digital assessment offers a feasible option because it uses simple tools while retaining authentic assessment principles.

At the same time, low-tech digital assessment should be understood as a contextual innovation rather than a technological limitation. In this study, low-tech digital refers to accessible tools such as Google Forms, Google Sheets, fillable rubrics, shared documents, or lightweight scoring templates that lecturers can use without complex infrastructure. These tools can help lecturers record scores, provide feedback, organize student submissions, and compare performance across criteria. Benchadlia et al. (2023) note that nursing simulation and educational technology may face barriers related to resources and teacher preparation. Salifu et al. (2022) also emphasize that nursing education in low-resource settings requires frameworks that fit local constraints while supporting clinical competence. Thus, low-tech digital design can reduce implementation barriers while keeping the assessment systematic and replicable.

Accordingly, the proposed assessment product adopts digital simplicity as a deliberate design choice. The goal is not to replace teacher judgment with automation but to support human scoring through clear digital templates and structured feedback fields. This is important because authentic nursing writing involves contextual judgment, clinical appropriateness, ethical tone, and genre sensitivity that automated tools may not interpret accurately. Astutik et al., (2024) caution that automated writing evaluation should be integrated critically because teachers still play an important role in interpreting feedback and maintaining pedagogical relevance. Hu et al. (2025) also emphasize that digital authentic assessment requires alignment among task design, technology use, and learning purposes. Therefore, the low-tech digital format in this study is used to enhance assessment practicality without weakening professional and pedagogical judgment.

However, these research streams have not been sufficiently integrated into a practical assessment product that combines authentic nursing writing tasks, analytic scoring criteria, expert validation, pilot testing, and accessible digital use. This gap is important because assessment shapes what lecturers teach, what students value, and how classroom writing practice prepares students for clinical communication. Therefore, this study addresses the lack of a validated and practically tested low-tech digital writing assessment for English for Nursing students. Its novelty lies in developing an assessment instrument with authentic nursing-related writing tasks, converting it into a simple digital format that can be used without advanced software or expensive platforms, and testing its validity and preliminary usability through expert judgment and limited classroom implementation

METHODS

1. Research Design

This study employed an Educational Research and Development design using a modified ADDIE model. The main output was a low-tech digital writing assessment package for English for Nursing students. The product included authentic writing tasks, an analytic scoring rubric, a digital scoring sheet, feedback forms, and lecturer guidelines. The development process consisted of analysis, design, development, expert validation, revision, limited implementation, and evaluation. Data were collected through both qualitative and quantitative procedures. Qualitative data came from document analysis, expert comments, lecturer feedback, and student responses, while quantitative data came from expert validation scores, practicality ratings, rubric usability scores, and student writing scores during pilot testing. This design allowed the researcher to examine the product's content relevance, clarity, usability, practicality, and preliminary classroom application.

2. Product Specification and Development Procedure

The product developed in this study was a low-tech digital authentic writing assessment package for English for Nursing students. The package consisted of five components: an authentic writing task

bank, an analytic scoring rubric, a digital scoring sheet, a feedback template, and a lecturer guideline. The task bank included five nursing-related writing tasks: patient admission note, nursing progress summary, patient education message, discharge instruction, and written SBAR summary. These tasks were selected because they represent writing activities that require clarity, accuracy, clinical vocabulary, professional tone, and patient-centered communication.

The analytic rubric assessed seven criteria: vocabulary, grammar, clinical accuracy, clarity, professional tone, empathy, and task completion. Each criterion used five performance levels to help raters evaluate students' writing consistently. The digital scoring sheet was designed with accessible low-tech tools so lecturers could record scores, organize results, and document feedback easily. The feedback template was linked to the rubric criteria, allowing students to receive specific comments on their strengths and areas for improvement. The lecturer guideline explained task administration, scoring procedures, feedback use, and score interpretation.

The development procedure followed a modified ADDIE sequence. The analysis stage identified course needs, nursing-related writing demands, and existing assessment practices. The design stage produced the assessment blueprint, writing prompts, rubric criteria, scoring scale, and digital format. The development stage created the first prototype of the task bank, rubric, scoring sheet, feedback template, and lecturer guideline. The validation stage involved expert review of content relevance, language clarity, clinical appropriateness, assessment quality, and digital usability. The product was revised based on expert comments and then implemented in a limited classroom pilot. The evaluation stage analyzed validation results, practicality data, user feedback, and pilot writing scores to refine the final product.

3. Research Instruments and Data Sources

This study used several instruments to collect data during the analysis, validation, and pilot testing stages. The first instrument was a document analysis checklist used to examine the English for Nursing syllabus, learning outcomes, teaching materials, writing tasks, and existing assessment practices. The second instrument was a need analysis form used to identify students' writing difficulties, nursing-related writing needs, rubric expectations, feedback needs, and digital access conditions. The third instrument was an expert validation sheet used to evaluate the product in terms of content relevance, task authenticity, clinical appropriateness, rubric clarity, language quality, scoring feasibility, feedback usefulness, and digital usability. During pilot testing, the study used a writing task sheet, an analytic rubric, a digital scoring sheet, practicality questionnaires, and an interview or feedback form. The writing task sheet guided students in completing authentic nursing-related writing tasks, while the analytic rubric assessed vocabulary, grammar, clinical accuracy, clarity, professional tone, empathy, and task completion. Practicality questionnaires collected student and lecturer responses on access, clarity, usefulness, time efficiency, and classroom feasibility. Interview or feedback forms provided qualitative comments for product revision.

4. Data Collection Procedure

Data collection was conducted through six procedural stages. First, the analysis phase reviewed the English for Nursing syllabus, learning outcomes, teaching materials, existing writing tasks, and current assessment practices. This phase also collected need analysis data from students and lecturer input on writing difficulties, nursing communication needs, rubric expectations, feedback needs, and digital access. Second, the prototype development stage produced the assessment blueprint, five authentic writing tasks, analytic rubric, digital scoring sheet, feedback template, and lecturer guideline. Third, the prototype was submitted to expert validators to assess content relevance, task authenticity, clinical appropriateness, rubric clarity, language quality, scoring feasibility, feedback usefulness, and digital usability. Fourth, the product was revised based on expert scores and written comments. Fifth, the revised product was implemented in a limited pilot test with English for Nursing students. Students completed authentic writing tasks, and the lecturer scored them using the analytic rubric and digital scoring sheet. Sixth, user feedback was collected through practicality questionnaires, interviews, or feedback forms to support final product refinement.

5. Data Analysis Technique

The data were analyzed using qualitative and quantitative techniques. Qualitative data from document analysis, expert comments, lecturer feedback, student responses, interviews, and feedback forms were analyzed through data condensation, categorization, interpretation, and conclusion drawing. The results were used to identify product strengths, weaknesses, and revision needs. Quantitative data from expert validation were analyzed using Aiken's V to examine the content validity of the assessment product. Reliability data were analyzed using Cronbach's alpha to measure the internal consistency of the rubric criteria and the intraclass correlation coefficient (ICC) to examine inter-rater reliability. Student writing scores from the pilot test were analyzed using descriptive statistics, including mean, standard deviation, minimum score, maximum score, and score distribution. Practicality questionnaire data were analyzed using percentage scores to determine the practicality level of the product from student and lecturer perspectives. When pre-test and post-test writing scores were available, a paired-sample t-test was used to examine whether there was a significant difference in students' writing performance after using the assessment product. The results of all analyses were integrated to determine the product's validity, reliability, practicality, usability, and preliminary effectiveness.

Results

This section presents the results of the Research and Development process for the low-tech digital writing assessment in English for Nursing. The findings are organized into six areas: need analysis, product development and revision, expert validity, reliability, practicality, and preliminary effectiveness. The limited pilot involved 37 English for Nursing students and three validators, consisting of one English language expert and two nursing education experts.

1. Need Analysis

The need analysis showed a strong demand for a more contextual writing assessment in English for Nursing. The overall mean score was 4.43, which indicated a very high level of need. Students reported that they needed clearer assessment standards, clinically relevant writing tasks, and feedback that could help them improve their nursing-related English writing. The need for a clear writing rubric obtained a mean score of 4.57, while the need for clinically relevant writing tasks reached 4.53. These findings show that students expected writing assessment to reflect nursing documentation, clinical communication, and professional language use.

A major challenge also appeared in the use of medical terminology. In total, 89.2% of students agreed or strongly agreed that they had difficulty using medical terms in written communication. This finding supports the inclusion of vocabulary, clinical accuracy, clarity, and professional tone in the assessment rubric. The students also viewed the low-tech digital format as feasible. Digital accessibility obtained a mean score of 4.30, and mobile phone access reached 4.46. Therefore, the product needed to combine authentic nursing writing tasks, analytic scoring, simple digital access, and structured feedback.

Table 1. Summary of Need Analysis Domains

Domain	Mean	Interpretation
Writing challenges	4.28	Very high
Clinical relevance of writing tasks	4.53	Very high
Need for analytic rubric	4.57	Very high
Need for feedback	4.38	Very high
Digital accessibility	4.38	Very high
Overall need analysis score	4.43	Very high

2. Product Development and Revision

The development phase produced a low-tech digital writing assessment package for English for Nursing students. The final product consisted of six components: an authentic writing task bank, an analytic scoring rubric, a digital scoring form, an automatic score recap sheet, a digital feedback

template, and a lecturer guideline. The task bank included nursing-related writing tasks such as patient admission notes, nursing progress summaries, patient education messages, discharge instructions, and written SBAR summaries. These tasks required students to write in clinically relevant contexts with appropriate vocabulary, organization, tone, and patient-centered language.

The analytic rubric used seven criteria: vocabulary, grammar, clinical accuracy, clarity, professional tone, empathy, and task completion. Each criterion used five performance levels to help lecturers assess writing more consistently. Expert feedback led to several revisions before pilot testing. The main revisions focused on descriptor wording, patient safety indicators, empathy, digital navigation, and clinical task prompts. The descriptors were rewritten into more observable statements. Patient safety elements were strengthened in clinical accuracy and task completion. The empathy criterion was revised to include respectful tone, reassurance, and sensitivity to patient condition. The digital format was also simplified to make scoring and feedback easier to use in class.

Table 2. Product Components and Revision Summary

Product Component	Description	Main Revision
Authentic writing task bank	Five nursing-related writing tasks, including admission note, progress summary, patient education message, discharge instruction, and written SBAR summary	Clinical prompts were clarified with patient background, writing purpose, and target audience
Analytic writing rubric	Seven criteria with five performance levels	Descriptors were rewritten into observable and measurable statements
Digital scoring form	Simple form for entering scores for each criterion	Navigation was simplified into scoring and feedback sections
Automatic score recap sheet	Spreadsheet-based summary for total scores and criterion averages	Score display was made clearer for lecturer use
Digital feedback template	Structured feedback linked to rubric criteria	Feedback wording was aligned with the seven rubric criteria
Lecturer guideline	Short guide for administration, scoring, and score interpretation	Instructions were clarified for classroom implementation

3. Expert Validity

Expert validity was examined through judgment from one English language expert and two nursing education experts. The validation used Aiken's V to measure the content validity of the product. The average Aiken's V was 0.90, which placed the product in the very valid category. The highest values appeared in the relevance of the assessment construct and the authenticity of nursing writing tasks, both with Aiken's V of 1.00. These results indicate that the validators strongly agreed that the product assessed the intended English for Nursing writing competence.

The lowest values appeared in instruction clarity, language quality, digital usability, and feedback format, each with Aiken's V of 0.83. These scores remained valid, but they also showed areas that needed refinement before pilot testing. The revisions focused on simplifying instructions, improving rubric language, clarifying feedback format, and making the digital layout easier to navigate. Overall, the validation results showed that the product was conceptually appropriate and ready for limited pilot testing after revision.

Table 3. Expert Validation Results Using Aiken's V

Validation Aspect	Validator Scores	Aiken's V	Interpretation
Relevance of assessment construct	5, 5, 5	1.00	Very valid
Authenticity of nursing writing tasks	5, 5, 5	1.00	Very valid
Alignment between task and rubric criteria	4, 5, 5	0.92	Very valid

Validation Aspect	Validator Scores	Aiken's V	Interpretation
Clinical accuracy of task content	4, 5, 5	0.92	Very valid
Clarity of writing task instructions	4, 5, 4	0.83	Valid
Language quality of rubric descriptors	5, 4, 4	0.83	Valid
Suitability of scoring procedure	5, 4, 5	0.92	Very valid
Usability of low-tech digital format	4, 4, 5	0.83	Valid
Usefulness of feedback template	4, 4, 5	0.83	Valid
Overall product feasibility	5, 5, 4	0.92	Very valid
Average	-	0.90	Very valid

4. Reliability Analysis

Reliability was examined through internal consistency and inter-rater agreement. The rubric produced a Cronbach's alpha of 0.891, which indicated high internal consistency among the seven criteria. This result shows that vocabulary, grammar, clinical accuracy, clarity, professional tone, empathy, and task completion worked together as a coherent assessment construct for nursing-related English writing.

Inter-rater reliability was also strong. Three raters scored the same 37 student texts using the revised rubric. The ICC single measure was 0.82, while the ICC average measure was 0.93. The exact agreement rate reached 76.0%, and the adjacent agreement rate reached 94.0%. This means that most scoring differences occurred within one score level. The findings suggest that the rubric could support consistent scoring when raters used the same criteria and descriptors.

Table 4. Inter-Rater Reliability of the Writing Assessment

Reliability Indicator	Result
Number of raters	3
Number of student texts	37
ICC single measure	0.82
ICC average measure	0.93
Exact agreement	76.0%
Adjacent agreement	94.0%
Interpretation	Strong inter-rater reliability

5. Practicality Analysis

The practicality analysis showed that the product was feasible for classroom use. The student practicality mean was 4.41, equivalent to 88.2%, which placed the product in the very practical category. Students gave the highest score to ease of access, with a mean of 4.54. This result indicates that they could access and complete the assessment through simple digital tools without advanced software. Feedback usefulness also received a high mean score of 4.49, showing that students valued the criterion-based feedback.

The lowest student score appeared in time efficiency, with a mean of 4.22. Although this score remained practical, it suggests that students needed more time to understand the clinical scenario, read the rubric, and complete the writing task carefully. Expert-user practicality also supported these findings. The average expert-user practicality score was 4.56, equivalent to 91.2%. The validators considered the lecturer guideline, digital scoring form, feedback template, score recap sheet, and classroom feasibility very practical. They recommended a short orientation before classroom use to help students understand the task format and rubric.

Table 5. Student and Expert-User Practicality Results

Group	Indicator or Component	Mean	Percentage	Interpretation
Student practicality	Ease of access	4.54	90.8%	Very practical

Group	Indicator or Component	Mean	Percentage	Interpretation
Student practicality	Digital navigation	4.38	87.6%	Very practical
Student practicality	Clarity of task instructions	4.43	88.6%	Very practical
Student practicality	Clarity of rubric criteria	4.35	87.0%	Very practical
Student practicality	Usefulness of feedback	4.49	89.8%	Very practical
Student practicality	Time efficiency	4.22	84.4%	Practical
Student practicality	Authenticity of writing task	4.46	89.2%	Very practical
Student practicality	Overall satisfaction	4.41	88.2%	Very practical
Student practicality	Average	4.41	88.2%	Very practical
Expert-user practicality	Lecturer guideline	4.67	93.4%	Very practical
Expert-user practicality	Digital scoring form	4.56	91.2%	Very practical
Expert-user practicality	Feedback template	4.44	88.8%	Very practical
Expert-user practicality	Score recap sheet	4.56	91.2%	Very practical
Expert-user practicality	Classroom feasibility	4.56	91.2%	Very practical
Expert-user practicality	Average	4.56	91.2%	Very practical

6. Preliminary Effectiveness

Preliminary effectiveness was examined by comparing students' writing performance before and after the use of the developed assessment package in the limited pilot. The pre-test required students to write a nursing-related text using a conventional instruction sheet. The post-test used a comparable task supported by the revised rubric, digital scoring format, and structured feedback. The writing scores were converted into a 0 to 100 scale.

The mean pre-test score was 68.84, while the mean post-test score increased to 79.27. The mean gain was 10.43 points, and the N-gain was 0.34, which indicated moderate improvement. The paired-sample t-test showed a significant difference between pre-test and post-test scores, $t(36) = 8.72$, $p < .001$. Cohen's d was 1.43, indicating a large practical effect in the limited pilot context. These results suggest that the product helped students understand writing expectations, use rubric-based feedback, and improve nursing-related written communication.

At the criterion level, improvement appeared across all rubric criteria. The largest improvements were found in clinical accuracy and professional tone. This pattern indicates that students became more aware of how to write clinically relevant and professionally appropriate texts after using the assessment package. However, these findings should be interpreted as preliminary because the study used one class, a limited number of participants, and no control group.

Table 6. Results of Preliminary Effectiveness from Pilot Testing

Component	Result
Pre-test mean	68.84
Pre-test SD	7.42
Post-test mean	79.27
Post-test SD	6.51
Mean gain	10.43
N-gain	0.34
N-gain interpretation	Moderate
95% confidence interval	8.02 to 12.84
t value	8.72
df	36
p value	< .001

Component	Result
Cohen's d	1.43
Interpretation	Significant preliminary improvement

Discussion

The findings indicate that the developed low-tech digital writing assessment addressed a specific gap in English for Nursing assessment. The need analysis showed that students required writing tasks that were closer to nursing documentation, patient-centered communication, medical terminology use, and professional written interaction. This result supports the principle of English for Specific Purposes, which places learner needs and target situation demands as the basis for course and assessment design (Hutchinson & Waters, 1987). In this study, the high need for clinical writing tasks and explicit rubrics suggests that students did not view writing as a general language activity only. They viewed writing as a professional skill connected to future nursing practice. This interpretation is consistent with previous studies on English for Nursing, which found that nursing students need English materials and assessments related to medical terminology, patient communication, written reports, and clinical documentation (Marleni et al., 2023; SB, 2023; Sudusinghe & Gamage, 2023). Therefore, the product contributes by transforming general writing assessment into contextual assessment that reflects the communicative demands of nursing education.

The expert validity and reliability results further show that the product had sufficient technical quality for limited classroom use. From the perspective of language assessment, a useful test or assessment should demonstrate construct validity, reliability, authenticity, practicality, and positive classroom impact (Bachman & Palmer, 1996). The expert validation results support construct relevance because the assessment measured writing through nursing-related tasks rather than isolated grammar or vocabulary items. The high validation scores for assessment construct and task authenticity indicate that the product matched the target domain of English for Nursing. The reliability findings also strengthen this interpretation. The rubric criteria worked consistently, and the inter-rater agreement showed that different raters could apply the rubric with strong consistency. This pattern supports the role of expert judgment in R&D because validation helps refine content, descriptors, format, and usability before implementation (Asli et al., 2024; Carey et al., 2022; Kania et al., 2024). Thus, the validation stage strengthened the academic credibility of the product and ensured that the assessment measured the intended writing construct.

The practicality findings show that low-tech digital assessment can offer a realistic alternative for English for Nursing classrooms with ordinary technological resources. This point is important because digital assessment is often associated with advanced platforms, paid systems, or complex automation. The product in this study used accessible tools such as a digital scoring form, spreadsheet-based recap sheet, structured feedback template, and lecturer guideline. This design supports Bachman and Palmer's (1996) concept of practicality, which emphasizes the relationship between assessment quality and available resources. The findings also align with Fitria (2023) study, which reported that Google Forms can support English assessment because it is easy to use, free, integrated with Google Sheets, and practical for lecturers and students. This finding is consistent with studies showing that digital assessment should be evaluated not only by technological sophistication but also by accessibility, usability, and pedagogical fit (Benchadlia et al., 2023; Hu et al., 2025).

The preliminary effectiveness findings suggest that the assessment product may support students' development in nursing-related English writing. The improvement from pre-test to post-test indicates that students performed better after using authentic tasks, explicit rubric criteria, digital scoring, and structured feedback. This finding can be explained through feedback theory. Hattie and Timperley (2007) argue that feedback becomes effective when it helps learners understand the goal, evaluate their current performance, and identify the next steps for improvement. This finding supports the argument that authentic assessment can guide learning when it clarifies expectations, connects tasks to professional practice, and provides actionable feedback (Nieminen et al., 2023; Vlachopoulos & Makri, 2024; Widanta et al., 2023). The result also supports Brookhart's (2018) argument that rubrics can make expectations visible and guide learning when criteria and performance levels are clearly

described. However, this finding should be interpreted as preliminary evidence only. The study used one group without a control group, so the improvement cannot be attributed solely to the assessment product.

This study has several limitations. The pilot involved only one class of 37 English for Nursing students, so the findings cannot be generalized to wider nursing education contexts. The study also did not use a control group, which limits the strength of the effectiveness claim. In addition, the implementation period was limited, so the study could not examine long-term writing development, repeated use of rubric-based feedback, or sustained improvement in specific writing criteria. The inter-rater reliability result was promising, but the study did not test systematic inter-rater training across a larger group of lecturers. Future research should involve larger samples, multiple nursing institutions, longer implementation periods, and comparison or control groups. The study contributes to ESP assessment by showing that a low-tech digital product can be valid, reliable, practical, and pedagogically useful when it is built from authentic nursing writing tasks and systematic R&D procedures. In this sense, the novelty of the study lies not in sophisticated technology but in the integration of authentic writing assessment, nursing documentation, analytic scoring, digital simplicity, and classroom feasibility (Benchadlia et al., 2023; Fisma, 2024; Hu et al., 2025).

Conclusion

This study shows that a low-tech digital writing assessment is relevant and feasible for English for Nursing students. The product connects writing assessment with real nursing communication needs, including clinical documentation, medical terminology, patient-centered expression, professional tone, and task completion. The findings indicate that the product has strong content validity, reliable scoring quality, high practicality, and promising preliminary impact on students' writing performance. Its main contribution lies in providing an accessible assessment model that can be used with simple digital tools and limited classroom resources. However, the results remain preliminary because the pilot involved one class, 37 students, and no control group. Future research should involve larger samples, multiple institutions, comparison groups, longer implementation, and systematic inter-rater training. Further studies should also examine how repeated rubric-based feedback improves clinical accuracy, empathy, and professional tone in nursing-related English writing.

References

- Almeman, K., EL Ayeb, F., Berrima, M., Issaoui, B., & Morsy, H. (2025). The Integration of AI and Metaverse in Education: A Systematic Literature Review. *Applied Sciences (Switzerland)*, 15(2), 50–65. <https://doi.org/10.3390/app15020863>
- Asli, N. F., Mohd Matore, M. E. E., & Md Yunus, M. (2024). Construct validity of primary trait writing rubrics based on assessment use argument (AUA) validation framework. *Heliyon*, 10(22), e40053. <https://doi.org/10.1016/j.heliyon.2024.e40053>
- Astutik, I., Widiati, U., Ratri, D. P., Jonathans, P. M., Nurkamilah, N., Devanti, Y. M., & Harfal, Z. (2024). Transformative Practices: Integrating Automated Writing Evaluation in Higher Education Writing Classrooms - A Systematic Review. *Indonesian Journal on Learning and Advanced Education (IJOLAE)*, 6(3), 423–441. <https://doi.org/10.23917/ijolae.v6i3.23675>
- Benchadlia, F., Rabia, Q., & Abderrahim, K. (2023). Simulation in Basic Nursing Student Education: Uses and Barriers. *The Open Nursing Journal*, 1–8. <https://doi.org/10.2174/18744346-v17-230717-2023-198>
- Brookhart, S. M. (2018). Appropriate Criteria: Key to Effective Rubrics. *Frontiers in Education*, 3(April). <https://doi.org/10.3389/feduc.2018.00022>
- Carey, M. D., Davidow, S., & Williams, P. (2022). Re-imagining narrative writing and assessment: a post-NAPLAN craft-based rubric for creative writing. *Australian Journal of Language and Literacy*, 45(1), 33–48. <https://doi.org/10.1007/s44020-022-00004-4>
- Fisma. (2024). Assessing the English Language Needs of Nursing Students: A Systematic Review. *JOEY: Journal of English Ibrahimy*, 3(2), 1–10. <https://doi.org/10.35316/joey.2024.v3i2.1-10>
- Fitria, T. N. (2023). Using Google Forms as an Online Learning Assessment Tool for Non-EFL Students. *Journal of English Language Teaching, Linguistics, and Literature Studies*, 3(2), 64. <https://doi.org/10.30984/jeltis.v3i2.2641>

- Harrison, N., Edmonds, M., Meads, C., Abdulmohdi, N., Prothero, L., & Shaw, S. (2024). *Simulation in nursing education: An evidence base for the future*. Council of Deans of Health and Anglia Ruskin University.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112. <https://doi.org/10.3102/003465430298487>
- Hu, A., Liu, Q., & Daniel, B. (2025). Digital Technologies in Authentic Assessment in Higher Education: A Systematic Literature Review and Narrative Synthesis. *SAGE Open*, 15(3). <https://doi.org/10.1177/21582440251357198>
- Huseynova, F. (2023). Authentic assessment in developing students' english language skills. *Bulletin of Postgraduate Education (Series)*, 25(54), 86–104. [https://doi.org/10.58442/2218-7650-2023-25\(54\)-86-104](https://doi.org/10.58442/2218-7650-2023-25(54)-86-104)
- Hutchinson, T., & Waters, A. (1987). *English for specific purposes: A learning-centred approach*. Cambridge University Press.
- Kania, N., Kusumah, Y. S., Dahlan, J. A., Nurlaelah, E., Gürbüz, F., & Bonyah, E. (2024). Constructing and providing content validity evidence through the Aiken's V index based on the experts' judgments of the instrument to measure mathematical problem-solving skills. *REID (Research and Evaluation in Education)*, 10(1), 64–79. <https://doi.org/10.21831/reid.v10i1.71032>
- Khojasteh, L., Kafipour, R., Pakdel, F., & Mukundan, J. (2025). Empowering medical students with AI writing co-pilots: design and validation of AI self-assessment toolkit. *BMC Medical Education*, 25(1). <https://doi.org/10.1186/s12909-025-06753-3>
- Marleni, L., Syarif, H., & Zainil, Y. (2023). ESP for Nurse: A Curriculum Analysis. *Journal of Education Research*, 4(4), 2452–2460. <https://doi.org/10.37985/jer.v4i4.757>
- Maxnun, L., Kristiani, K., & Sulistyaningrum, C. D. (2024). Development of hots-based cognitive assessment instruments: ADDIE model. *Journal of Education and Learning*, 18(2), 489–498. <https://doi.org/10.11591/edulearn.v18i2.21079>
- Nieminen, J. H., Bearman, M., & Ajjawi, R. (2023). Designing the digital in authentic assessment: is it fit for purpose? *Assessment and Evaluation in Higher Education*, 48(4), 529–543. <https://doi.org/10.1080/02602938.2022.2089627>
- Orfanò, B. M., & Wingate, U. (2024). Teaching Academic Writing To Nursing Students: Developing Criticality Through a Genre-Based Approach. *ESP Today*, 12(1), 118–135. <https://doi.org/10.18485/esptoday.2024.12.1.6>
- Rohmani, N., Zulkarnaen, D., & Winar Cahyo, P. (2024). Web-Based Nursing Care Documentation for Students to Support Online Learning. *International Review of Research in Open and Distributed Learning*, 25(4), 52–65. <https://doi.org/10.19173/irrodl.v25i4.7758>
- Salifu, D. A., Heymans, Y., & Christmals, C. Dela. (2022). A Simulation-Based Clinical Nursing Education Framework for a Low-Resource Setting: A Multimethod Study. *Healthcare (Switzerland)*, 10(9), 1–23. <https://doi.org/10.3390/healthcare10091639>
- Salifu, D. A., Heymans, Y., & Christmals, C. Dela. (2022). Facilitating the Development of Clinical Competence in a Low-Resource Setting: Perceptions and Challenges of Nurse Educators. *Nurse Media Journal of Nursing*, 12(1), 42–60. <https://doi.org/10.14710/nmjn.v12i1.43995>
- Sari, T. S. (2026). *English For Nursing: A Need Analysis of Contextual- Based Language Teaching*. 15(1), 1283–1292.
- SB, Handayani. (2023). Enhancing English For Nursing: A Comprehensive Needs Analysis For Effective Communication In Emergency Nursing Study Programs. *Journal of Humanity Studies*, 2(2), 107–124. <https://doi.org/10.22202/jhs.2023.v2i2.7686>
- Sekarsari, A. A., Muamaroh, M., & Hikmat, M. H. (2023). Authentic Assessment of Speaking Skills in Junior High School. *IDEAS: Journal on English Language Teaching and Learning, Linguistics and Literature*, 11(1), 630–637. <https://doi.org/10.24256/ideas.v11i1.3854>
- Setiawan, A., Cendana, W., Ayres, M., Yuldashev, A. A., & Setyawati, S. P. (2023). Development and validation of a self-assessment-based instrument to measure elementary school students' attitudes in online learning. *REID (Research and Evaluation in Education)*, 9(2), 184–197. <https://doi.org/10.21831/reid.v9i2.52083>

- Sudusinghe, W. S., & Gamage, C. K. W. (2023). The Importance of English Language Proficiency for the Success of Nursing Education in Sri Lanka. *South-East Asian Journal of Medical Education*, 17(1), 39–42. <https://doi.org/10.4038/seajme.v17i1.529>
- Vlachopoulos, D., & Makri, A. (2024). A systematic literature review on authentic assessment in higher education: Best practices for the development of 21st century skills, and policy considerations. *Studies in Educational Evaluation*, 83(November), 101425. <https://doi.org/10.1016/j.stueduc.2024.101425>
- Widanta, I. M. R. J., Putra, I. M. A., Ardika, I. W. D., & Hidayanti, N. N. A. T. (2023). Developing English for Nurses in Indonesia: From Learning Supporting Tools to Assessment. *Theory and Practice in Language Studies*, 13(4), 917–926. <https://doi.org/10.17507/tpls.1304.13>
- Zaim, M., Refnaldi, & Arsyad, S. (2020). Authentic assessment for speaking skills: Problem and solution for english secondary school teachers in Indonesia. *International Journal of Instruction*, 13(3), 587–604. <https://doi.org/10.29333/iji.2020.13340a>