Journal Of Nursing Practice
https://thejnp.org/
ISSN: 2614-3488 (print); 2614-3496 (online)
Vol. 10 No.1 October. Page 263-281
https://doi.org/10.30994/jnp.v10i1.1289

Developing the ForCAP Framework for Competency Assessment in Care Plan Formulation: A Delphi Study and Preliminary Psychometric Evaluation

Mohamad As'ad Efendy[1*], Michael I. Aggari[2]
[1] *Faculty of Nursing, Universitas Strada Indonesia, Kediri, Indonesia*
[2] *School of Nursing, Philippine Women's University, Manila, Philippines*
***Corresponding author:** surel.asad@gmail.com*

ORCID: **Mohamad A. Efendy:** https://orcid.org/0000-0003-2736-7593

ABSTRACT

Background: Care plan formulation is a core nursing competency, but instruments that target it are limited, particularly in Indonesia.

Purpose: To develop the Framework for Competency Assessment in Care Plan Formulation (ForCAP), to examine expert consensus on its structure, and to obtain preliminary content-validity, reliability, and dimensionality evidence for an aligned student self-assessment questionnaire.

Methods: This sequential multiphase instrument-development study reviewed 16 competency frameworks (2010–2023), conducted a two-round modified Delphi with 15 purposively selected experts (acceptance: at least 70% rating 3 or 4), and had the same experts rate the content validity of a 32-item questionnaire, which was then piloted with 224 nursing students.

Results: The initial 57 elements were refined to 5 domains, 10 competencies, and 32 sub-competencies; all 42 final elements were accepted (Fleiss' kappa 0.93). Item-level content validity ranged from 0.80 to 0.93 (scale-level 0.82–0.86). Total-score Cronbach's alpha was 0.87 (95% CI 0.85–0.89); domain alphas were 0.46–0.68.

Conclusion: Experts endorsed the ForCAP structure, and the questionnaire total score has preliminary support. The faculty-rated rubric, proficiency anchors, and progression indicators remain provisional and need further validation.

Keywords : Clinical competence; Delphi technique; Education, Nursing; Nursing Process; Psychometrics

BACKGROUND

Nursing education has shifted toward competency-based education (CBE) so that graduates have the practical skills that healthcare demands (Gravina, 2017; Lewis et al., 2022). Formulating a comprehensive care plan is central to this shift: it operationalizes patient-centred care and supports the quality and continuity of nursing care (American Association of Colleges of Nursing [AACN], 2021; Ballantyne, 2016). Competence is understood as an integration of

The Journal of Nursing Practice, its website, and the articles published there in are licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

Received: October 13, 2026; Revised: October 26, 2026; Accepted: October 27, 2026; Available Online: October 28, 2026
Copyright © [2026] Author.

knowledge, skills, attitudes, and clinical judgement that develops progressively through practice (Fukada, 2018; Nabizadeh-Gharghozar et al., 2021).

Existing instruments address adjacent constructs rather than care-plan formulation across the nursing process (Table 1). Recent work includes the Nursing Student Competence Scale (30 items, six factors; Huang et al., 2022), the Competency of Nursing Process Questionnaire for registered nurses in Cambodia (24 items, five factors corresponding to assessment, diagnosis, planning, implementation, and evaluation; Koy et al., 2023), the Clinical Reasoning Competency Scale for nurses (Bae et al., 2023), a simulation-based care assessment rubric with content validation only (Pang & Yuan, 2026), and the Student Survey on Writing Nursing Care Plans with its Persian and Turkish adaptations (Aldemir Atmaca et al., 2025; Ramazanzadeh et al., 2023; Salvador et al., 2022). General frameworks such as the Competency Outcomes and Performance Assessment model capture broad clinical performance rather than care planning (Lenburg, 1999; Redman et al., 1999). Reviews describe heterogeneous tools with incomplete psychometric evidence and assessors who struggle to interpret competency language (Alkhelaiwi et al., 2024; Immonen et al., 2019). Without a dedicated tool, programs may rely on instructor-dependent judgements, and students who excel in theory may struggle to turn it into structured care plans, reflecting the enduring theory–practice gap (Saifan et al., 2021).

Table 1. Selected instruments related to care-plan competence

Instrument	Construct and rater	Evidence reported by developers	Relation to ForCAP
Q-DIO (Šerková & Marečková, 2018)	Quality of documented nursing diagnoses, interventions, and outcomes; rater of patient records	Expert consensus for content validity; intra- and inter-rater $\kappa = 0.95$; $\alpha = 0.83–0.99$ by concept	Rates documentation, not student competence
Lasater Clinical Judgment Rubric (Lasater, 2007)	Clinical judgment in simulation; evaluator rubric	Developed from simulation observation; later reliability studies exist	Clinical judgment, not care planning
Student Survey on Writing Nursing Care Plans (Salvador et al., 2022; adapted by Ramazanzadeh et al., 2023; Aldemir Atmaca et al., 2025).	Students' care-plan writing along the nursing process; student survey	Mixed-methods development (22 interviews, 195 students); Persian version reliable and convergently valid; Turkish version (n = 362) $\alpha = 0.98$	Closest comparator; no SDKI/SIKI/SLKI terminology or proficiency levels
Competency of Nursing Process Questionnaire (Koy et al., 2023)	Nursing-process competency of registered nurses; self-report	24 items; I-CVI 1.00; $\alpha = 0.963$; test–retest 0.769; five factors (70.1% variance); n = 260	Practising nurses, not students; developed outside Indonesia

Vol. 10 No.1 October. Page 263-281
https://doi.org/10.30994/jnp.v10i1.1289

Instrument	Construct and rater	Evidence reported by developers	Relation to ForCAP
Nursing Student Competence Scale (Huang et al., 2022)	General competence of baccalaureate students; self-report	30 items, six factors; acceptable CFA fit; α = 0.91–0.98; test–retest 0.52 (n = 30); n = 241	General competence, not care planning
Clinical Reasoning Competency Scale (Bae et al., 2023)	Clinical reasoning of hospital nurses; self-report	22 items, three factors; α = 0.92; Delphi panel during development; n = 483	Clinical reasoning of nurses, not care-plan formulation
Comprehensive Nursing Care Assessment Rubric (Pang & Yuan, 2026)	Nursing care competencies in simulation; observer rubric	Content validation only (five experts; S-CVI/Ave = 0.99)	Simulation; focus on physical assessment and clinical judgment
ForCAP (this study)	Care-planning competence across the nursing process; student self-assessment (faculty-rated version untested)	Delphi consensus (n = 15); I-CVI; α, ω, EFA (N = 224)	Adds expert-consensus structure, Indonesian terminologies, proficiency descriptors; psychometric evidence is preliminary

Indonesia has standardized nursing languages for diagnoses, interventions, and outcomes (SDKI, SIKI, SLKI; PPNI, 2016, 2018a, 2018b), yet their integration into curricula is under-explored, and almost half of the nurses in one hospital study did not use standardized diagnoses in documentation (Somantri et al., 2021). Clinical training is also constrained by limited instructor availability, feedback, and simulation facilities (Naryati et al., 2025), and global frameworks such as the AACN Essentials were not designed around SDKI, SIKI, and SLKI (AACN, 2021).

The study was situated within three perspectives: Benner's (1984) novice-to-expert model, which provided the staged proficiency scale (Murray et al., 2019), and Bruner's (1966) constructivism for scaffolded, performance-based learning and Watson's (1985) theory of human caring for relational and ethical behaviours, which were used afterwards to interpret the framework. These perspectives were not used as an analytic coding scheme. *Construct definition:* in this study, care-planning competence refers to the knowledge, skills, and judgement needed to formulate, implement, and evaluate a nursing care plan across the nursing process, from assessment to evaluation. "Care plan formulation" in the framework name therefore denotes this plan-centred competence, including enactment and evaluation of the plan.

Among the instruments we identified through a non-systematic search, none combines an expert-consensus structure derived from national competency frameworks, the Indonesian

Vol. 10 No.1 October. Page 263-281
https://doi.org/10.30994/jnp.v10i1.1289

standardized terminologies, and proficiency descriptors for formative self-assessment of student care planning; this statement does not claim priority. ForCAP was designed to fill this gap. The objective was to develop ForCAP and to obtain preliminary content-validity, internal-consistency, and dimensionality evidence for the aligned questionnaire. The research questions were: (1) What initial framework can be developed from an annotated review of existing competency frameworks? (2) What are the level and strength of expert agreement on the initial framework? (3) What final framework results from refinement, and what are the level and strength of expert agreement on it? (4) What provisional framework, including proficiency scale and assessment criteria, can be proposed for further testing? (5) What preliminary content-validity, internal-consistency, and dimensionality evidence supports the aligned student self-assessment questionnaire?

METHODS

Design and Setting

We conducted a sequential multiphase instrument-development study (Boateng et al., 2018; Creswell & Plano Clark, 2017) integrating document analysis, Delphi ratings and comments, and questionnaire psychometric testing (Figure 1). Mixed-methods reporting (GRAMMS; O'Cathain et al., 2008) was applied because qualitative expert comments and quantitative ratings were integrated to decide the fate of each element (joint display, Table 5), and Delphi reporting followed CREDES (Jünger et al., 2017). Data were collected between December 2024 and January 2025, after ethical approval. Pilot testing took place at one Indonesian university, across its undergraduate nursing and Nursing Education (Pendidikan Ners) programs.

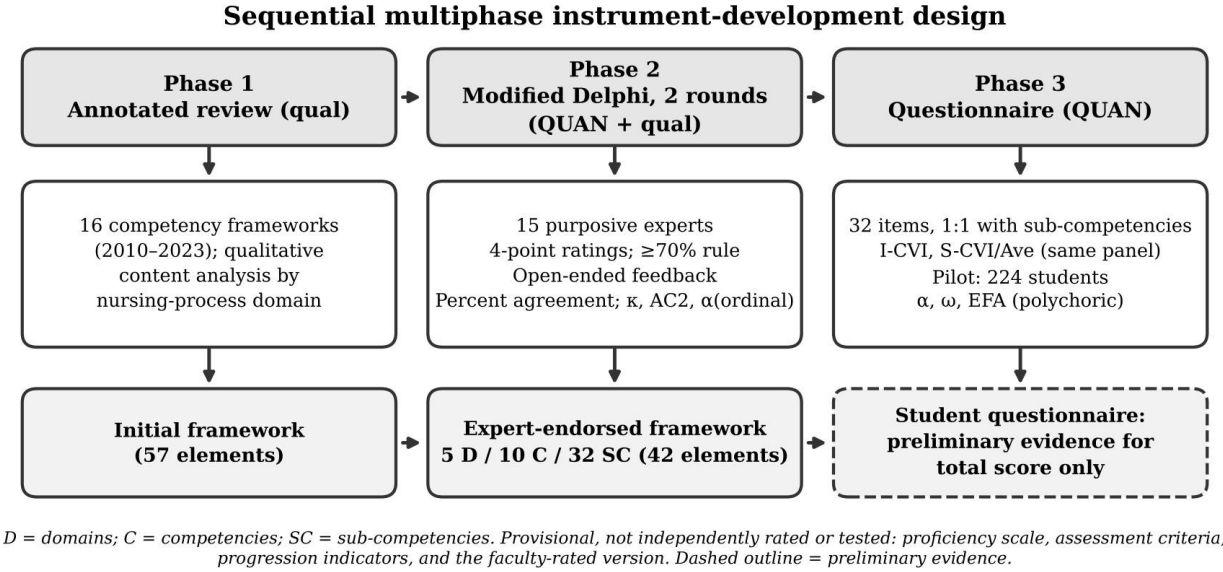


Figure 1. Sequential multiphase design, integration points, and status of the outputs

Participants and Sampling

Phase 1 (documents). Competency frameworks were purposively identified by navigating the official websites of nursing regulatory bodies, health ministries, and professional organizations. Documents were eligible if they (a) were issued by a national

Vol. 10 No.1 October. Page 263-281
https://doi.org/10.30994/jnp.v10i1.1289

regulatory authority or recognized nursing organization, (b) specified core competencies for registered, entry-to-practice, or nurse practitioner roles relevant to the nursing process, (c) were published between 2010 and 2023, a window chosen to capture frameworks issued after the global shift toward competency-based nursing education and to keep the sources current for contemporary practice, and (d) were accessible in English or Indonesian. Sixteen frameworks (ASEAN, 10; North America, 4; Australia, 1; Europe, 1) were included. Three nurse practitioner frameworks (CNA, CNO, NONPF) and one general-practice standard were retained because they describe assessment, diagnostic reasoning, and care-planning behaviours in detail; they informed sub-competency wording, whereas advanced-practice content (e.g., prescribing) was not adopted.

Phase 2 (experts). A purposive panel of 15 experts was recruited through professional networks, academic institutions, and nursing associations; all 15 invited experts participated in both rounds. Inclusion criteria were at least 5 years of experience in nursing education or practice involving curriculum development, competency-based assessment, or clinical training; expertise in care planning, clinical reasoning, or competency assessment; and at least a master's degree. Small, homogeneous panels are considered adequate when members hold relevant expertise (Nasa et al., 2021; Niederberger & Spranger, 2020).

Phase 3 (students). Students were recruited by purposive sampling from one CBE-oriented university (undergraduate nursing and Ners programs); faculty members helped identify eligible students but took no part in obtaining consent or collecting data; consent was obtained by the research team. Eligible participants were final-year baccalaureate students, students in the professional (Ners) program, or recent Ners graduates, all with documented experience developing nursing care plans. The minimum sample was set with Cochran's (1977) formula (95% confidence, p = 0.50, margin of error 0.05, accessible population of 500), giving 218. Because this formula targets a proportion rather than factor-analytic power, adequacy for preliminary factor analysis was judged by overall sample size (N = 224; 200–300 is commonly recommended; Boateng et al., 2018), the respondent-to-item ratio (224/32 = 7.0, below the 10:1 often advised), and empirical factorability (KMO, Bartlett's test).

Instruments and Data Collection

Three instruments were used; all were developed and administered in Indonesian. First, the Delphi rating form listed each competency and sub-competency for rating (the five nursing-process domains were fixed a priori and not rated) on a four-point scale: 4 = strongly agree (highly appropriate, no change needed), 3 = agree (appropriate, minor changes), 2 = disagree (major changes needed), 1 = strongly disagree (delete or replace entirely). Ratings thus combine agreement with the amount of revision needed, not the importance of content; 3 or 4 means acceptance with at most minor changes. Experts could also comment on each domain. Second, the content validity form rated each questionnaire item for relevance, clarity, and alignment with its sub-competency on a four-point scale (Lynn, 1986; Polit & Beck, 2006). Third, the ForCAP self-assessment questionnaire comprised 32 first-person behavioural statements mapped one-to-one onto the sub-competencies of the final framework. Items were written directly from the sub-competencies, without a separate item-generation stage or cognitive interviews, and several statements combine more than one behaviour. Items are rated

Vol. 10 No.1 October. Page 263-281
https://doi.org/10.30994/jnp.v10i1.1289

1–5 on the proficiency levels adapted from Benner (1 = novice to 5 = expert), each with a brief first-person descriptor; level 0 (absent) was omitted from the student questionnaire because all respondents had documented care-plan experience, so a response indicating no performance of the behaviour was considered conceptually inapplicable, and to keep the scale to five categories. Because Benner's stages describe development through experience rather than response categories, this response format is provisional. Total and domain scores are means (range 1–5), interpreted descriptively; no cut-off scores were derived. The English item wording in this article was prepared by the authors and has not been back-translated or validated.

Phase 1 used a structured annotation template (publisher, scope, care-planning competencies, limitations). In Phase 2, both rounds were administered online (Google Forms; two-week window), with open-ended prompts to justify ratings and to propose merging, splitting, rewording, relocating, or removing elements; before the final round experts received item-level agreement percentages and anonymized comments. All 15 experts completed both rounds. In Phase 3, the same experts rated the questionnaire items, after which the questionnaire was administered online to students. The first author conducted the Delphi process and coding under the second author's supervision; coding was reviewed by the second author without independent double coding or member checking. The first author held no supervisory role over student participants.

Data Analysis

Delphi. Agreement per element was the proportion of experts rating 3 or 4. The ≥70% threshold (Diamond et al., 2014) was an a priori decision rule for retention and is not itself evidence of construct validity; qualitative feedback then guided merging, restructuring, relocating, or rewording. Fleiss' κ (Fleiss, 1971; Landis & Koch, 1977) treats ratings as nominal and is sensitive to skewed distributions, so Gwet's AC2 with quadratic weights (Gwet, 2008) and Krippendorff's ordinal α (Hayes & Krippendorff, 2007) were also computed, together with κ restricted to non-unanimous elements and exact (Clopper–Pearson) 95% confidence intervals for proportions. Qualitative content analysis (Phases 1–2) was coded by the first author and reviewed by the second; discrepancies were resolved by discussion. Findings were integrated with a joint display classifying each pairing as convergent, divergent, or complementary.

Content validity. I-CVI (proportion of 15 experts rating 3 or 4; criterion 0.78), modified kappa (k*; Polit et al., 2007), S-CVI/Ave (≥0.80 acceptable, ≥0.90 excellent), and S-CVI/UA were computed for each criterion (Lynn, 1986; Polit & Beck, 2006; Yusoff, 2019). The full panel is the primary analysis; a reduced-panel analysis was added only as an exploratory sensitivity analysis.

Pilot data. Reliability was evaluated with Cronbach's α (percentile bootstrap 95% CI, 5,000 resamples; ≥0.70 acceptable; Taber, 2018), McDonald's ω (one-factor maximum-likelihood model), mean inter-item correlation, corrected item–total correlations, and domain α. The response options of Item 10 (prioritizing diagnoses) were displayed in reverse order in the online form; the item was recoded (6 − x) in the primary analysis, and results are also reported as recorded and with Item 10 excluded. Dimensionality was examined with KMO,

Bartlett's test, permutation parallel analysis (95th percentile, 1,000 permutations), and exploratory factor analysis (minimum-residual extraction) on polychoric correlations because responses were ordinal, comparing one-factor and five-factor solutions with oblique (oblimin) and orthogonal (varimax) rotation.

Ethical Considerations

Ethical approval was granted by the Health Research Ethics Committee of Universitas Strada Indonesia (Approval No. 001911/EC/KEPK/I/12/2024, issued in 2024) before data collection. All experts and students gave written informed consent after full information on purpose, procedures, and the right to withdraw without penalty. Students were told that the pilot evaluated the instrument, not their individual competency. Recruitment and data collection were conducted independently of academic assessment; participation was voluntary, responses were anonymous, and neither participation nor responses affected grades or progression. Data were anonymized, stored in password-protected files, and accessible only to the research team. No identifiable information or images are reported, so consent for publication was not applicable. No clinical intervention was involved, so trial registration was not applicable.

RESULTS

Participant Characteristics

The expert panel (N = 15) was mostly aged 35–39 years (47%), female (60%), and highly qualified (73% Master of Nursing, 27% Doctor of Nursing), with balanced academic (53%) and healthcare (47%) settings. Students (N = 224) were mostly aged 23–25 years (69%) and female (72%); 69% had training in clinical documentation standards (Table 2).

Table 2. Characteristics of the expert panel and student participants

Characteristic	n (%)
Expert panel (N = 15)	
Age: 30–34 / 35–39 / >40 years	2 (13.3) / 7 (46.7) / 6 (40.0)
Gender: female / male	9 (60.0) / 6 (40.0)
Qualification: Master / Doctor of Nursing	11 (73.3) / 4 (26.7)
Setting: academic / healthcare institution	8 (53.3) / 7 (46.7)
Position: administrator / faculty / nurse manager / clinical nurse	4 (26.7) / 4 (26.7) / 3 (20.0) / 4 (26.7)
Experience: 5–10 / 11–15 / >16 years	2 (13.3) / 8 (53.3) / 5 (33.3)
Expertise†: curriculum/standards / competency-based assessment / supervision	10 (66.7) / 9 (60.0) / 5 (33.3)
Students (N = 224)	
Age: 20–22 / 23–25 / >25 years	64 (28.6) / 154 (68.8) / 6 (2.7)
Gender: female / male	162 (72.3) / 62 (27.7)
Status: final-year bachelor / Ners program / Ners graduate	71 (31.7) / 92 (41.1) / 61 (27.2)

Vol. 10 No.1 October. Page 263-281
https://doi.org/10.30994/jnp.v10i1.1289

Characteristic	n (%)
Training: documentation standards / NANDA-I, NOC, NIC (NNN) and SDKI/SIKI/SLKI / electronic health records	154 (68.8) / 83 (37.1) / 52 (23.2)

Initial Framework and Expert Feedback

Content analysis of the 16 frameworks under the five nursing-process domains yielded an initial framework of 57 elements (15 competencies, 42 sub-competencies). Expert comments converged on four themes: (a) merging overlapping items (e.g., holistic assessment with tools) while splitting physical examination and interview techniques; (b) mandating NANDA International (NANDA-I)/SDKI, NOC/SLKI, and NIC/SIKI; (c) strengthening client/family engagement and cultural considerations; and (d) relocating health education to planning and embedding quality and emergency elements elsewhere. Illustrative comments were “Must include NANDA/SDKI as standard classification” (Expert 2) and “Education fits better in planning” (Expert 9). Removed clusters were described as “Too broad for general framework” (Expert 4).

Expert Agreement on the Initial and Final Frameworks

In Round 1, 37 of 57 elements met the 70% rule; 20 elements were rated 1 by all 15 experts and removed: five clusters (health risks, healthcare needs, resources, emergencies, quality; 5 competencies and 13 sub-competencies) and two further sub-competencies (under C3.2 and C5.2). One sub-competency of the healthcare-needs cluster was retained at 73.3% and merged into C2.1. Because comments mostly recommended integrating this content elsewhere, these ratings more plausibly reflect disagreement with standalone inclusion than with the content’s importance, which the form could not test. Retained elements with 73–80% agreement (11 elements) drew ratings of 2 from three or four experts each. After refinement, 42 elements (10 competencies and 32 sub-competencies, i.e., the 27 retained sub-competencies plus five added from expert comments; Table 5) were re-rated: all met the criterion, with every expert rating each element 3 (31.7% of 630 ratings) or 4 (68.3%) (Table 3). Agreement coefficients were high ($\kappa = 0.783$ and 0.934; AC2 = 0.941 and 0.995; ordinal $\alpha = 0.929$ and 0.934), but they were driven by unanimity: 46 of 57 Round 1 elements and 33 of 42 final-round elements were rated identically by all 15 experts. Among the elements on which experts differed, agreement was at chance level in Round 1 ($\kappa = -0.04$; mean inter-expert correlation −0.05) and substantial in the final round ($\kappa = 0.65$). With 15 experts, 100% agreement has an exact 95% CI of 78.2%–100%. Because the same panel rated revisions it had proposed, final-round agreement confirms acceptance of the revisions and is not independent validation.

Table 3. Expert ratings and agreement indices by Delphi round

Indicator	Round 1 (initial)	Final round
Elements rated (competencies + sub-competencies)	57 (15 + 42)	42 (10 + 32)
Ratings of 1 / 2 / 3 / 4, %	35.1 / 4.4 / 8.3 / 52.2 (of 855)	0 / 0 / 31.7 / 68.3 (of 630)

Indicator	Round 1 (initial)	Final round
Elements with ≥70% rating 3 or 4	37 (20 removed)	42
Elements rated identically by all 15 experts	46 (20 all 1; 26 all 4)	33 (11 all 3; 22 all 4)
Fleiss' κ, all elements	0.783	0.934
Fleiss' κ, non-unanimous elements only	−0.04 (11 elements)	0.65 (9 elements)
Gwet's AC2 (quadratic weights)	0.941	0.995
Krippendorff's α (ordinal)	0.929	0.934

Validity and Reliability of the Questionnaire

All 32 items exceeded the I-CVI criterion on relevance, clarity, and alignment (0.80–0.93; relevance 0.80 for 25 items, 0.87 for 4, and 0.93 for 3; exact 95% CI for 12/15 = 52%–96%), with S-CVI/Ave of 0.82, 0.86, and 0.86 (acceptable, below 0.90) and k* of 0.80–0.93 (Table 4). No item was rated 3 or 4 by every expert (S-CVI/UA = 0 for all three criteria). Three experts (E3, E7, E11) gave 186 of 221 ratings below 3 (84.2%) and rated 20%, 27%, and 59% of all ratings 3 or 4, against 96–99% for each of the other 12 experts. These ratings are lower than the unanimous Delphi endorsement of the same sub-competencies; they concern a different object (suitability of a first-person statement versus importance of the competency), and the experts' reasons were not recorded, so their objections could not be analysed. In an exploratory analysis without these three experts, S-CVI/Ave was 0.91 (relevance) and 1.00 (clarity, alignment); this is descriptive and does not replace the full-panel result. The same panel produced the Delphi and content-validity ratings.

Pilot data from 224 students had no missing responses. Item means ranged from 2.12 to 3.84, option 3 accounted for 42% of responses, and the total mean score was 3.01 (SD = 0.37; range 2.25–3.91), with no floor or ceiling effect. For the total score, α was 0.869 (95% CI 0.848–0.885), ω was 0.877, and the mean inter-item correlation was 0.18 (0.23 on polychoric correlations). Domain α values were 0.46 (Assessment), 0.49 (Diagnosis), 0.68 (Planning), 0.52 (Implementation), and 0.60 (Evaluation), all below 0.70, so domain subscores are not reliable. Corrected item–total correlations were 0.25–0.48; only Items 1 (0.25) and 25 (0.28) were below 0.30. Item 10 correlated −0.37 with the total score as recorded and 0.37 after recoding; total α was 0.850 as recorded, 0.869 after recoding, and 0.866 with Item 10 excluded (Domain 2 α: 0.15 as recorded, 0.49 after recoding). Items 20, 23, and 24 had bimodal response distributions (SD > 1.2).

Sampling adequacy was acceptable (KMO = 0.75; Bartlett's $\chi^2(496) = 1{,}649.9$, $p < .001$). In permutation parallel analysis the first eigenvalue (6.73) far exceeded the random-data threshold (1.89), whereas the second (1.76) was indistinguishable from its threshold (1.75–1.76), indicating one dominant factor. On polychoric correlations, the first eigenvalue was 8.37 (26.1% of variance), and a one-factor model had loadings of 0.34–0.60 and 23.8% explained variance, but a residual RMSR of 0.076, so unidimensionality is not established. Five-factor oblimin and varimax solutions did not reproduce the five a priori domains (items from the same domain loaded on different factors; factor correlations 0.10–0.39).

Journal Of Nursing Practice
https://thejnp.org/
ISSN: 2614-3488 (print); 2614-3496 (online)
Vol. 10 No.1 October. Page 263-281
https://doi.org/10.30994/jnp.v10i1.1289

Table 4. Content validity and reliability of the ForCAP questionnaire (32 items)

Property	Value	Interpretation
I-CVI (relevance, clarity, alignment)	0.80–0.93	≥0.78; all items retained
S-CVI/Ave (relevance; clarity; alignment)	0.82; 0.86; 0.86	≥0.80 acceptable; <0.90
S-CVI/UA (all three criteria)	0.00	No item rated 3–4 by all 15 experts
Modified kappa, k* (relevance, clarity, alignment)	0.80–0.93	>0.74 excellent (Polit et al., 2007)
Cronbach's α, total (95% CI)	0.869 (0.848–0.885)	Item 10 recoded; 0.850 as recorded; 0.866 without Item 10
McDonald's ω (one factor)	0.877	—
Mean inter-item r	0.18	Low
Domain α (D1–D5)	0.46; 0.49; 0.68; 0.52; 0.60	All <0.70; subscores not supported
Corrected item–total r	0.25–0.48	Items 1 and 25 <0.30
KMO; parallel analysis	0.75; 1 dominant factor	Five domains not reproduced

Integration of Findings and Recommended Framework

The strands did not diverge (Table 5; codes are Round 1 codes unless marked final). For removed clusters, 0% agreement converged with comments describing them as too broad or better handled elsewhere. For retained components, agreement showed that a component belonged in the framework while comments specified its restructuring; health education, for example, had 100% agreement yet was relocated from implementation to planning.

Table 5. Joint display integrating Round 1 agreement and expert feedback

Round 1 component	Agreement	Illustrative expert feedback	Integration and decision
C1.1 Holistic assessment, tools, examination and interview	73–80%	"Merge for comprehensive approach" (E7); "Physical examination and interview techniques should be separate" (E1)	Complementary; tools merged, examination and interview split
C1.3 Health risks; C2.2 healthcare needs; C3.3 resources; C4.4 emergencies; C5.3 quality	0%	"Too broad" (E4); "Better suited for emergency-specific competencies" (E7); "Quality aspects should be integrated throughout" (E14)	Convergent; removed as standalone, content integrated elsewhere
C2.1 Nursing diagnoses	80%	"Must include NANDA/SDKI" (E2)	Complementary; retained sub-competency of C2.2 merged in; NANDA-I/SDKI added

Round 1 component	Agreement	Illustrative expert feedback	Integration and decision
C3.1 Care plan; C3.2 client-centred planning	73%	"Include NOC/SLKI and NIC/SIKI" (E5); "Add cultural considerations" (E13)	Complementary; terminologies, cultural and continuity elements added
C4.1 Interventions	73%	"Specify safety protocols" (E3); "Add technical skills standards" (E15)	Complementary; safety, evidence, skills elements added
C4.3 Health education	100%	"Education fits better in planning" (E9)	Complementary; relocated to planning (final C3.3)
C5.1 Care effectiveness	80%	"Add standardized criteria" (E4)	Complementary; NOC/SLKI-based evaluation

The expert-endorsed structure is shown in Table 6, and the status of every ForCAP component in Table 7. The proposed ForCAP couples this structure with a six-point proficiency scale adapted from Benner (0 = absent, 1 = novice, 2 = advanced beginner, 3 = competent, 4 = proficient, 5 = expert; the questionnaire uses levels 1–5), assessment criteria for the 32 sub-competencies, and progression indicators. The criteria and indicators were drafted by the first author, were not rated by the experts or tested for inter-rater reliability, and are provisional.

Table 6. Expert-endorsed ForCAP structure (5 domains, 10 competencies, 32 sub-competencies)

Domain / competency	Sub-competencies
D1. Assessment	
C1.1 Conducts comprehensive health assessment	Performs comprehensive assessment using appropriate tools and techniques; conducts systematic physical examination; demonstrates effective interview techniques
C1.2 Analyses assessment data	Analyses clinical data to identify patterns and variations; applies critical thinking; integrates evidence-based practice
D2. Diagnosis	
C2.1 Formulates and prioritizes healthcare problems	Analyses assessment data to identify actual/potential problems; uses diagnostic reasoning to validate problems; formulates diagnoses with a standardized system (NANDA-I/SDKI); prioritizes diagnoses by clinical significance; involves client, family, and team
D3. Planning	
C3.1 Develops care plan	Formulates plan with measurable outcomes and interventions; establishes priorities by urgency and risk; plans outcomes using NOC/SLKI; plans interventions using NIC/SIKI
C3.2 Ensures client-centred planning	Engages client/family in decision-making; maintains care continuity across settings; incorporates cultural considerations
C3.3 Plans health education	Develops targeted education strategies; designs self-care management plans; plans teaching evaluation methods

Domain / competency	Sub-competencies
D4. Implementation	
C4.1 Executes nursing interventions	Implements evidence-based interventions following safety protocols; maintains therapeutic environment and nurse–client relationship; performs technical nursing skills according to standards (NIC/SIKI)
C4.2 Manages dynamic situations	Responds to changing client conditions; modifies interventions based on response; coordinates interdisciplinary care
D5. Evaluation	
C5.1 Evaluates care outcomes	Monitors and evaluates outcomes using standardized criteria (NOC/SLKI); validates outcomes through systematic assessment; documents findings accurately and in a timely manner
C5.2 Modifies care plan	Revises interventions based on evaluation findings; updates expected outcomes based on client response

NANDA-I = NANDA International (diagnoses); NIC = Nursing Interventions Classification; NOC = Nursing Outcomes Classification; SDKI/SIKI/SLKI = Indonesian standards for diagnoses/interventions/outcomes.

Table 7. Validation status of ForCAP components

Component	Evidence in this study	Status
Domains, competencies, sub-competencies	Annotated review; two-round Delphi, 15 experts	Expert-endorsed
Student questionnaire (32 items)	I-CVI/S-CVI; α, ω, EFA (N = 224)	Preliminary; total score only
Proficiency scale (0–5; questionnaire uses 1–5)	Adapted from Benner; not rated	Provisional
Assessment criteria; progression indicators	Drafted by first author; not rated	Provisional
Faculty-rated version	Not tested	Untested

DISCUSSION

This study developed an expert-endorsed framework for care-planning competence and obtained preliminary psychometric evidence for an aligned student questionnaire, addressing a gap in specialized tools (Redman et al., 1999; Sanchez, 2019; Immonen et al., 2019). The Delphi process indicated agreement on the domain–competency–sub-competency structure, whereas faculty-rated assessment anchors, inter-rater reliability, and criterion validity remain untested, and the alignment with Bruner's and Watson's theories is interpretive rather than an analytical validation result. Compared with recent instruments that reported test–retest reliability and factor-analytic evidence (Huang et al., 2022; Koy et al., 2023), the psychometric evidence for ForCAP is thinner, and its distinctive contribution lies in the expert-consensus structure, the Indonesian terminologies, and the proficiency descriptors.

First, the consolidation from 57 to 42 elements reflects a move from exhaustive lists toward an integrated model. Experts did not endorse standalone competencies for risk evaluation, healthcare needs, resource management, emergency response, and quality, but their comments favoured embedding this content rather than judging it unimportant. This resonates

Vol. 10 No.1 October. Page 263-281
https://doi.org/10.30994/jnp.v10i1.1289

with Salvador et al. (2022), who organized students' care-plan writing difficulties along the five steps of the nursing process. The rise in κ from 0.783 to 0.934 should be read as acceptance of revisions by the panel that proposed them; ordinal α hardly changed (0.929 to 0.934), so the apparent gain in κ should not be over-interpreted. The high agreement also needs cautious reading: most elements received identical ratings from all experts, and such uniformity may reflect shared interpretation of the rating question, consensus pressure, or acquiescence in iterative designs (Barrett & Heale, 2020; Niederberger & Spranger, 2020); replication with an independent panel is needed.

Second, in the final round all 15 experts rated the elements incorporating standardized classification systems (NANDA-I/SDKI, NIC/SIKI, NOC/SLKI) 3 or 4. ForCAP translates SDKI/SIKI/SLKI (PPNI, 2016, 2018a, 2018b) into assessable statements, which responds to evidence that standardized languages improve documentation and communication (Adubi et al., 2018; Enebeli et al., 2023; Silva et al., 2022; Zhang et al., 2021; Emmanuel et al., 2024) and to inconsistent use among Indonesian nurses (Somantri et al., 2021); whether ForCAP improves such outcomes is untested. Third, coupling the nursing process with Benner's (1984) progression offers a developmental alternative to pass/fail judgement (Murray et al., 2019). The inclusion of client/family engagement and cultural considerations (Senarathne & Meegoda, 2021) is consistent with Watson's caring perspective, and the guided-to-independent level descriptors with Bruner's scaffolding; both mappings were made after the framework was developed and are interpretive.

Fourth, the questionnaire supports cautious use of the total score as a formative, self-reflective tool. Content validity was acceptable but not excellent (S-CVI/Ave 0.82–0.86). A total α of 0.869 in a heterogeneous sample indicates internal consistency, not construct validity (Taber, 2018): the mean inter-item correlation was low, domain α values were 0.46–0.68, and the data supported one dominant factor, so the five domains are a conceptual blueprint rather than an empirically supported score structure and domain subscores should not be used. Self-assessment can foster reflective practice (Andrade, 2019), but the response format, the item wording, and the self-report bias remain open issues.

The weak domain-level reliability and the absent five-factor structure deserve explicit interpretation. Cronbach's α depends on the number of items and their mean inter-item correlation (Taber, 2018); with only 5–10 items per domain and a mean inter-item correlation of 0.18, domain α values below 0.70 are expected, and α reached 0.869 only when all 32 items were pooled. Several features of the instrument may have contributed: items were written directly from sub-competencies, some combine more than one behaviour, and the Benner-based response format may be read as self-rated frequency by some students and as developmental stage by others (42% of responses were at the midpoint). The nursing process is also iterative and interdependent, so students may not experience assessment, diagnosis, planning, implementation, and evaluation as separable skills, and a single general factor may reflect an overall self-appraisal tendency rather than competence itself. Related instruments developed in other populations reported higher internal consistency and multi-factor structures (Huang et al., 2022; Koy et al., 2023); differences in item pools, respondents, and sample composition prevent direct comparison, but the present results show that a domain structure

derived from experts' conceptual consensus does not automatically translate into empirical subscales. Domain subscores should therefore not be used until items have been revised through cognitive interviews and the structure has been tested in an independent sample.

Limitations

The document sample was purposive rather than systematically searched, Europe is represented by one framework, and nurse practitioner and general-practice frameworks informed an entry-level framework. Sampling was purposive from one university, the respondent-to-item ratio was 7.0, 61 of 224 participants were graduates, and educational status was not linked to item responses, so subgroup and known-groups analyses were not possible. The study did not test the five-factor structure by confirmatory analysis, test–retest or inter-rater reliability, or criterion validity, and the assessment criteria, progression indicators, and faculty version were not validated. The response options of Item 10 were displayed in reverse order and the item was recoded post hoc; sensitivity analyses are reported. Domains 4 and 5 extend beyond plan formulation to implementation and evaluation, so the framework should be read as covering care planning across the nursing process.

Implications for Nursing Practice

ForCAP may guide formative discussion and faculty calibration about care-planning expectations, and the questionnaire may support self-reflection. Faculty-rated assessments should not be used for high-stakes decisions before rubric validation and inter-rater reliability testing. Use in classroom, simulation, and practicum settings, longitudinal tracking, comparison of perceived and observed competence, and benefits for employers are prospective applications that have not been evaluated.

CONCLUSION

This study developed ForCAP and provided expert-consensus evidence for its structure of 5 domains, 10 competencies, and 32 sub-competencies, together with preliminary psychometric evidence for an aligned student self-assessment questionnaire (S-CVI/Ave 0.82–0.86 with S-CVI/UA = 0; total-score $\alpha = 0.869$). The five-domain structure is conceptual: domain subscores were not reliable ($\alpha = 0.46$–0.68) and the data supported one dominant factor. The faculty-rated rubric, proficiency anchors, and progression indicators have not been validated. Future work should revise weak items and the response format using cognitive interviews, then pursue cross-institutional validation, confirmatory factor analysis, test–retest and inter-rater reliability, and criterion validity.

ACKNOWLEDGMENTS

The authors thank the expert panellists and student participants for their contributions.

AUTHOR CONTRIBUTIONS

Efendy was responsible for the conceptualization, methodology, investigation, formal analysis, and writing of the original draft. Aggari provided supervision, validation, and contributed to writing, reviewing, and editing. All authors have read and agreed to the published version of the manuscript, and agree to be personally accountable for their own contributions and the integrity of the work.

FUNDING

This research received no funding.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

DECLARATION OF AI USE

During manuscript preparation, the authors used Quillbot for grammar and language editing. The authors reviewed and edited the output and remain responsible for the content and integrity of the article.

REFERENCES

Adubi, I. O., Olaogun, A. A., & Adejumo, P. O. (2018). Effect of standardized nursing language continuing education programme on nurses' documentation of care at University College Hospital, Ibadan. *Nursing Open, 5*(1), 37–44. https://doi.org/10.1002/nop2.108

Aldemir Atmaca, K., Hançer Tok, H., & Yıldız, E. (2025). Turkish adaptation of the student survey on writing nursing care plans: A validity and reliability study. *BMC Nursing, 24*(1). https://doi.org/10.1186/s12912-025-04037-1

Alkhelaiwi, W. A., Traynor, M., Rogers, K., & Wilson, I. (2024). Assessing the competence of nursing students in clinical practice: The clinical preceptors' perspective. *Healthcare, 12*(10), 1031. https://doi.org/10.3390/healthcare12101031

American Association of Colleges of Nursing. (2021). *The essentials: Core competencies for professional nursing education*. Retrieved from https://www.aacnnursing.org/Portals/0/PDFs/Publications/Essentials-2021.pdf

Andrade, H. L. (2019). A critical review of research on student self-assessment. *Frontiers in Education, 4*, 87. https://doi.org/10.3389/feduc.2019.00087

Bae, J., Lee, J., Choi, M., Jang, Y., Park, C. G., & Lee, Y. J. (2023). Development of the clinical reasoning competency scale for nurses. *BMC Nursing, 22*, 138. https://doi.org/10.1186/s12912-023-01244-6

Ballantyne, H. (2016). Developing nursing care plans. *Nursing Standard, 30*(26), 51–57. https://doi.org/10.7748/ns.30.26.51.s47

Barrett, D., & Heale, R. (2020). What are Delphi studies? *Evidence-Based Nursing, 23*(3), 68–69. https://doi.org/10.1136/ebnurs-2020-103303

Benner, P. (1984). *From novice to expert: Excellence and power in clinical nursing practice*. Menlo Park, CA: Addison-Wesley.

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health, 6*, 149. https://doi.org/10.3389/fpubh.2018.00149

Bruner, J. S. (1966). *Toward a theory of instruction*. Cambridge, MA: Harvard University Press.

Cochran, W. G. (1977). *Sampling techniques* (3rd ed.). New York, NY: John Wiley & Sons.

Creswell, J. W., & Plano Clark, V. L. (2017). *Designing and conducting mixed methods research* (3rd ed.). Thousand Oaks, CA: SAGE Publications.

Diamond, I. R., Grant, R. C., Feldman, B. M., Pencharz, P. B., Ling, S. C., Moore, A. M., & Wales, P. W. (2014). Defining consensus: A systematic review recommends methodologic criteria for reporting of Delphi studies. *Journal of Clinical Epidemiology, 67*(4), 401–409. https://doi.org/10.1016/j.jclinepi.2013.12.002

Emmanuel, C. E., Blessing, O. J., & Paulina, A. A. (2024). Perception and utilization of standardized nursing languages among nurses in south-south Nigeria. *International Journal of Nursing and Midwifery, 16*(1), 1–8. https://doi.org/10.5897/ijnm2023.0536

Enebeli, E. C., Akpan-Idiok, P. A., & Chukwudozie, C. C. (2023). Nurses' lived experiences in the utilization of standardized nursing languages in documentation of nursing care, in Cross River State, Nigeria. *International Journal of Nursing Knowledge, 34*(4), 276–283. https://doi.org/10.1111/2047-3095.12400

Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin, 76*(5), 378–382. https://doi.org/10.1037/h0031619

Fukada, M. (2018). Nursing competency: Definition, structure and development. *Yonago Acta Medica, 61*(1), 1–7. https://doi.org/10.33160/yam.2018.03.001

Gravina, E. W. (2017). Competency-based education and its effect on nursing education: A literature review. *Teaching and Learning in Nursing, 12*(2), 117–121. https://doi.org/10.1016/j.teln.2016.11.004

Gwet, K. L. (2008). Computing inter-rater reliability and its variance in the presence of high agreement. British Journal of Mathematical and Statistical Psychology, 61(1), 29–48. https://doi.org/10.1348/000711006X126600

Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. Communication Methods and Measures, 1(1), 77–89. https://doi.org/10.1080/19312450709336664

Huang, S.-M., Fang, S.-C., Hung, C.-T., & Chen, Y.-H. (2022). Psychometric evaluation of a nursing competence assessment tool among nursing students: A development and validation study. *BMC Medical Education, 22*, 372. https://doi.org/10.1186/s12909-022-03439-y

Immonen, K., Oikarainen, A., Tomietto, M., Kääriäinen, M., Tuomikoski, A.-M., Kaučič, B. M., … Mikkonen, K. (2019). Assessment of nursing students' competence in clinical practice: A systematic review of reviews. *International Journal of Nursing Studies, 100*, 103414. https://doi.org/10.1016/j.ijnurstu.2019.103414

Jünger, S., Payne, S. A., Brine, J., Radbruch, L., & Brearley, S. G. (2017). Guidance on Conducting and REporting DElphi Studies (CREDES) in palliative care: Recommendations based on a methodological systematic review. *Palliative Medicine, 31*(8), 684–706. https://doi.org/10.1177/0269216317690685

Koy, V., Yunibhand, J., Rauth, A., Bircher, N., Prak, M., & Henker, R. (2023). Development and psychometric testing of a competency of nursing process questionnaire. *International Journal of Nursing Sciences, 10*(2), 245–250. https://doi.org/10.1016/j.ijnss.2023.03.009

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159–174. https://doi.org/10.2307/2529310

Lasater, K. (2007). Clinical judgment development: Using simulation to create an assessment rubric. Journal of Nursing Education, 46(11), 496–503. https://doi.org/10.3928/01484834-20071101-04

Lenburg, C. B. (1999). The framework, concepts and methods of the competency outcomes and performance assessment (COPA) model. *OJIN: The Online Journal of Issues in Nursing, 4*(2). https://doi.org/10.3912/OJIN.Vol4No02Man02

Lewis, L. S., Rebeschi, L. M., & Hunt, E. (2022). Nursing education practice update 2022: Competency-based education in nursing. *SAGE Open Nursing, 8*. https://doi.org/10.1177/23779608221140774

Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research, 35*(6), 382–385. https://doi.org/10.1097/00006199-198611000-00017

Murray, M., Sundin, D., & Cope, V. (2019). Benner's model and Duchscher's theory: Providing the framework for understanding new graduate nurses' transition to practice. *Nurse Education in Practice, 34*, 199–203. https://doi.org/10.1016/j.nepr.2018.12.003

Nabizadeh-Gharghozar, Z., Alavi, N. M., & Ajorpaz, N. M. (2021). Clinical competence in nursing: A hybrid concept analysis. *Nurse Education Today, 97*, 104728. https://doi.org/10.1016/j.nedt.2020.104728

Naryati, N., Hadi, M., Agung, R. N., & Fajarini, M. (2025). Challenges in clinical training for professional nurse students: A qualitative study. Jurnal Keperawatan Padjadjaran, 13(3), 281–292. https://doi.org/10.24198/jkp.v13i3.2700

Nasa, P., Jain, R., & Juneja, D. (2021). Delphi methodology in healthcare research: How to decide its appropriateness. *World Journal of Methodology, 11*(4), 116–129. https://doi.org/10.5662/wjm.v11.i4.116

Niederberger, M., & Spranger, J. (2020). Delphi technique in health sciences: A map. *Frontiers in Public Health, 8*, 457. https://doi.org/10.3389/fpubh.2020.00457

O'Cathain, A., Murphy, E., & Nicholl, J. (2008). The quality of mixed methods studies in health services research. *Journal of Health Services Research & Policy, 13*(2), 92–98. https://doi.org/10.1258/jhsrp.2007.007074

Pang, W. I. P., & Yuan, H. (2026). Development and content validation of a comprehensive nursing care assessment rubric (CNAR) in simulation. *Teaching and Learning in Nursing*. https://doi.org/10.1016/j.teln.2026.06.011

Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health, 29*(5), 489–497. https://doi.org/10.1002/nur.20147

Polit, D. F., Beck, C. T., & Owen, S. V. (2007). Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. Research in Nursing & Health, 30(4), 459–467. https://doi.org/10.1002/nur.20199

PPNI. (2016). *Standar diagnosis keperawatan Indonesia: Definisi dan indikator diagnostik* (1st ed.). Jakarta, Indonesia: DPP PPNI.

PPNI. (2018a). *Standar intervensi keperawatan Indonesia: Definisi dan tindakan keperawatan* (1st ed.). Jakarta, Indonesia: DPP PPNI.

PPNI. (2018b). *Standar luaran keperawatan Indonesia: Definisi dan kriteria hasil keperawatan* (1st ed.). Jakarta, Indonesia: DPP PPNI.

Ramazanzadeh, N., Ghahramanian, A., Zamanzadeh, V., Onyeka, T. C., Valizadeh, L., & Ghaffarifar, S. (2023). Translation, cross-cultural adaptation and evaluation of the Iranian Persian version of student survey on writing nursing care plan. *Nursing Open, 10*(8), 5476–5484. https://doi.org/10.1002/nop2.1786

Redman, R., Lenburg, C. B., & Walker, P. H. (1999). Competency assessment: Methods for development and implementation in nursing education. *OJIN: The Online Journal of Issues in Nursing, 4*(2). https://doi.org/10.3912/OJIN.Vol4No02Man03

Saifan, A., Devadas, B., Daradkeh, F., Abdel-Fattah, H., Aljabery, M., & Michael, L. M. (2021). Solutions to bridge the theory-practice gap in nursing education in the UAE: A qualitative study. *BMC Medical Education, 21*, 490. https://doi.org/10.1186/s12909-021-02919-x

Salvador, J. T., Alqahtani, F. M., Sauce, B. R. J., Alvarez, M. O. C., Rosario, A. B., Reyes, L. D. V., Mohamed, E. R., Awadh, L. A., Sanchez, K. K. B., Alzaid, M., Agman, D. D., & Schonewille, M. A. P. (2022). Development of student survey on writing nursing care plan: An exploratory sequential mixed-methods study. Journal of Nursing Management, 30(5). https://doi.org/10.1111/jonm.12996

Sanchez, C. L. (2019). Nursing care plan: An evidence-based tool for learning and providing high-quality care. *Journal of Nursing Education and Practice, 9*(6), 1–7. https://doi.org/10.5430/jnep.v9n6p1

Senarathne, H. S., & Meegoda, M. (2021). Comparative analysis of cultural competence of senior and junior nursing students. *Sri Lanka Journal of Health Research, 1*(1), 14–19. https://doi.org/10.4038/sljhr.v1i1.28

Šerková, D., & Marečková, J. (2018). Creation and validation of Q-DIO, an instrument for rating the quality of nursing documentation – literature review. Central European Journal of Nursing and Midwifery, 9(1), 799–810. https://doi.org/10.15452/CEJNM.2018.09.0007

Silva, M. B. da, Engelman, B., Figueiredo, M. S., & Almeida, M. de A. (2022). Standardized nursing languages in orthopedics from the perspective of the nursing role effectiveness model. *Revista Gaúcha de Enfermagem, 43*. https://doi.org/10.1590/1983-1447.2022.20210321.en

Somantri, I., Sekarwana, N., Mediawati, A. S., Ramoo, V., Yahaya, N. A., & Yudianto, K. (2021). Analysis of standardized nursing languages application in the context of nursing diagnosis in hospital in Indonesia. *Open Access Macedonian Journal of Medical Sciences, 9*(T6), 173–178. https://doi.org/10.3889/oamjms.2021.7307

Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education, 48*(6), 1273–1296. https://doi.org/10.1007/s11165-016-9602-2

Watson, J. (1985). *Nursing: The philosophy and science of caring*. Boulder, CO: University Press of Colorado.

Yusoff, M. S. B. (2019). ABC of content validation and content validity index calculation. *Education in Medicine Journal*, *11*(2), 49–54. https://doi.org/10.21315/eimj2019.11.2.6

Zhang, T., Wu, X., Peng, G., Zhang, Q., Chen, L., Cai, Z., & Ou, H. (2021). Effectiveness of standardized nursing terminologies for nursing practice and healthcare outcomes: A systematic review. *International Journal of Nursing Knowledge*, *32*(4), 220–228. https://doi.org/10.1111/2047-3095.12315