Molecular & Cellular Biomechanics 2025, 22(3), 1165.
https://doi.org/10.62617/mcb1165

Article

Development and reliability testing of the four-dimensional job performance evaluation questionnaire for clinical nurses in the biological perspective

Haoyu Chen[1], Ping Li[1,*], Ming Hong[2], Xiaohan Zhao[1], Tianqing Zhuang[1]

[1] Medical School, Shihezi University, Shihezi 832003, China
[2] Nursing Department, Xinjiang Uygur Autonomous Region People's Hospital, Urumqi 830054, China
*** Corresponding author:** Ping Li, 1450291938@qq.com

CITATION

Chen H, Li P, Hong M, et al. Development and reliability testing of the four-dimensional job performance evaluation questionnaire for clinical nurses in the biological perspective. Molecular & Cellular Biomechanics. 2025; 22(3): 1165. https://doi.org/10.62617/mcb1165

ARTICLE INFO

Received: 18 December 2024
Accepted: 8 January 2025
Available online: 12 February 2025

Abstract: Purpose: This study aimed to scientifically develop the Four-Dimensional Job Performance Evaluation Questionnaire for Clinical Nurses from a biological aspect and to assess its reliability and validity. **Methods:** Guided by Koopmans' four-dimensional work performance theory, the questionnaire was structured around task performance, relational performance, adaptive performance, and counterproductive performance. The initial version was developed through a comprehensive literature review, analysis of hospital performance evaluation indices, and Delphi method consultations with experts. A survey was conducted among 549 clinical nurses in a tertiary hospital in Xinjiang Province, and the questionnaire's reliability and validity were evaluated using the critical ratio method, factor analysis, and reliability and validity tests. **Results:** The content validity index (I-CVI) for each item ranged from 0.820 to 1.000, with an average level content validity index (S-CVI/Ave) of 0.970. The overall Cronbach's alpha coefficient for the questionnaire was 0.962, with dimension-specific coefficients of 0.967, 0.901, 0.953, and 0.909. Exploratory factor analysis indicated a cumulative variance contribution rate of 63.235% for the four principal factors, and confirmatory factor analysis confirmed a good model fit, leading to the finalization of a 45-item questionnaire covering four dimensions of clinical nurses' job performance. **Conclusion:** The Four-Dimensional Job Performance Evaluation Questionnaire for Clinical Nurses developed in this study demonstrates good reliability and validity, offering a comprehensive measurement tool for assessing nurses' job performance levels in the biological context.

Keywords: nurses; questionnaire development; job performance; adaptive performance; reliability; validity; biological perspective

1. Introduction

Nursing is a critical component of healthcare workforce resources [1]. Nursing personnel constitute over 40% of the health technician workforce, and their roles encompass all facets of patient care. The performance of nurses, which encompasses the behavioral outcomes and achievements associated with their nursing duties, is intrinsically linked to patient health [2]. "As public health demands escalate, nurse performance has emerged as a critical metric for assessing nursing staff and the quality of nursing services. References such as 'The National Nursing Development Plan (2021–2025)' [3] and the 'Notice of the General Office of the National Health Commission on Further Strengthening Nursing in Medical Institutions' [4] emphasize the need for hospitals to prioritize the innovation of nurse performance management models. These documents advocate for the establishment of a scientific and systematic performance evaluation system to motivate nurses and foster

continuous improvement and quality enhancement in nursing services. Building on this foundation, an increasing number of scholars are directing their attention to the study of nurses' work performance. Understanding the current state of nurses' work performance is essential for effectively managing nursing staff and developing targeted interventions to enhance their performance levels. Preliminary literature analysis revealed that existing research tools for assessing nurses' work performance [5], predominantly developed in earlier years, may not align with the current nursing work environment. These tools often lack a robust scientific theoretical foundation, utilize generic work performance surveys, and show limited relevance to the specific demands of nursing work, indicating a need for refinement. This study, grounded in scientific performance management theory, constructs a four-dimensional work performance evaluation questionnaire for clinical nurses by integrating the index system research of nursing performance evaluation in China. It aligns with the current specific requirements of nursing work and refines the questionnaire through the Delphi method. The study also conducts reliability testing to provide a measurement tool for a comprehensive understanding of nurses' work performance. Adhering to the Delphi Research Implementation and Reporting Criteria, this study reports the research process and outcomes [6]. Nursing work not only involves psychological and social support for patients, but is also closely related to the biological dimension. Nurses play a key role in maintaining the physiological stability of patients, which is closely related to the biological processes of the human body [7,8]. For example, nurses' monitoring of patients' vital signs and nursing interventions directly affect the physiological balance of the human body, just like the regulatory mechanism of environmental homeostasis in organisms. In addition, nursing work is closely related to the fields of human immunobiology and tissue repair biology. By preventing infection and promoting wound healing, the performance of nurses is crucial to the recovery of patients. Therefore, when evaluating the performance of nurses, biological factors must be taken into account to ensure the comprehensiveness and scientificity of the evaluation.

2. Methods

2.1. Initial questionnaire item pool development

2.1.1. Development of an entry pool from literature and nurse performance appraisal programs

The initial dimensions and item pool for the questionnaire were derived from Koopmans' four-dimensional work performance structure theory [9], supplemented by a literature analysis. By reviewing existing research on the work performance of nurses in China and assessing the limitations of current measurement tools [10–17], we integrated insights from the performance appraisal index systems of nurses in China and performance evaluation documents from three tertiary hospitals in the Xinjiang region. This process led to the initial construction of a four-dimensional work performance questionnaire for clinical nurses, encompassing a total of 54 items across task performance, relational performance, adaptive performance, and counterproductive performance.

2.1.2. Expert group discussion on item pool

Ten nursing management experts from a tertiary hospital in Shihezi City participated in an expert focus group discussion based on the literature-analyzed entry pool. This discussion led to the formation of an initial questionnaire pool comprising 49 items across four dimensions: task performance (16 items), relationship performance (11 items), adaptive performance (14 items), and counterproductive performance (8 items).

Task performance included items on nurses' mastery of professional skills, adherence to rules and regulations for clinical nursing work, health education, preventive healthcare, and continuity of nursing services. In addition, the application of biological knowledge is covered, i.e., nurses understand and correctly apply pharmacokinetics during drug administration (e.g., adjusting the dose of the drug according to the patient's liver and kidney function) to ensure the efficacy and safety of the drug. When dealing with genetic test results, nurses need to understand how specific genetic variants can lead to dysfunctional proteins or dysregulated signaling pathways that can lead to disease. Nurses should take appropriate nursing measures, such as personalized treatment protocols, to intervene in these pathological processes at the molecular level. At the same time, it also includes the biological monitoring of patients, including the accuracy and timeliness of the collection of biological samples (e.g., blood, urine), as well as the correct interpretation and recording of biological test results (e.g., blood glucose, blood lipids), these tasks reflect the performance of nurses in biological testing. Nurses are required to collect and preliminarily process biological samples, including blood, tissue, and other bodily fluids, during care. By analyzing these samples at the molecular level, subtle changes that have occurred in the patient's body can be revealed. For example, specific protein markers (e.g., C-reactive protein, tumor markers, etc.) and gene expression profiles (e.g., changes in mRNA levels) can be important indicators for early diagnosis and monitoring of diseases. Nurses need to be able to interpret the results of these molecular tests and adjust their care plans accordingly. In addition, the relationship between nursing operations and the biological structure and function of the human body is also addressed, for example, when performing venipuncture, the nurse's familiarity with the anatomy of human blood vessels and how to avoid damage to surrounding tissues and nerves, ensuring biologically safe operations.

Relationship performance encompassed voluntary behaviors and tasks unrelated to direct patient care, such as management, teaching, research, and team-building activities, as well as their contribution to hospital work effectiveness. In addition, it also covers the biological communication ability of nurses in interdisciplinary collaboration. When working with multidisciplinary teams such as doctors and nutritionists, nurses need to have certain biological knowledge for effective communication. For example, when discussing the nutritional support plan for patients, nurses should understand the biological processes of the body's digestion, absorption and utilization of nutrients (involving digestive system biology, cell metabolism biology, etc.) in order to better coordinate the team to provide comprehensive care for patients. When nurses provide health education to patients, they often explain a lot of biological knowledge, such as the pathogenesis of the

disease (explaining the cause from a biological perspective), the biological principles of treatment plans (such as the mechanism of action of chemotherapy drugs on cancer cells), etc. Assessing nurses' abilities and effectiveness in this regard can be part of relationship performance, which helps improve patients' understanding of their own diseases and their degree of cooperation with treatment, thereby affecting the treatment effect.

Adaptive performance measured the degree to which nurses adapt to changes in the nursing work environment or roles. It mainly includes the following two aspects: (1) The ability to cope with the development of biomedical technology: With the continuous advancement of biomedical technology, such as the application of new technologies such as genetic testing and biological therapy in clinical practice, nurses need to have the ability to quickly learn and adapt to these new technologies. For example, understand the impact of genetic test results on patient care plans (involving knowledge such as genetics and molecular biology) and how to integrate them into the nursing requirements of biological therapy (such as the observation and care of adverse reactions after immunotherapy, involving immune system biology). This requires nurses not only to master new technologies and tools, but also to be able to apply this knowledge to actual nursing work to ensure the quality and safety of patient care; (2) the ability to adapt to individual biological differences in patients: Each patient has individual differences in biology, such as genetic background, physiological function, etc. Nurses need to be able to adjust nursing strategies according to these biological characteristics of patients. For example, special care for allergic patients (involving individual differences in immune biology), personalized care for patients of different age groups (considering the biological characteristics of different growth and development stages), and other nursing adjustments based on individual differences in patients. This helps nurses provide more accurate and effective nursing services, thereby improving patients' treatment effects and nursing satisfaction. Counterproductive performance referred to behaviors contrary to nursing work goals, potentially harming nursing work and the work environment. Specifically, they include: (1) Violation of aseptic operation principles: For example, failure to comply with aseptic operation principles (based on the biological principles of microbiology and infection control) will increase the risk of infection for patients, thus seriously affecting the quality of care. Such behaviors can become important indicators of counterproductive performance, including but not limited to: failure to comply with aseptic technical operation specifications, leading to cross-infection or nosocomial infection; neglecting basic infection control measures such as hand hygiene, increasing the risk of pathogen transmission; and mishandling medical waste, which may lead to environmental pollution and pathogen spread; (2) behaviors that ignore the biological needs of patients: For example, failure to provide appropriate dietary guidance based on the patient's dietary restrictions (based on the patient's condition and biological metabolic needs), or failure to create a quiet and comfortable environment during the patient's rest and recovery period (following the human body clock and physiological recovery mechanism). These behaviors violate the biological requirements of nursing, may have a negative impact on the patient's recovery, and should be included in the evaluation of counterproductive performance. The questionnaire used a 5-point Likert scale, with task performance,

relationship performance, and adaptive performance scored from 1 (‘not at all’) to 5 (‘very much’), and counterproductive performance scored inversely from 5 (‘very much’) to 1 (‘not at all’).

2.2. Expert delphi consultation

The Delphi expert correspondence method was employed to refine and delete items from the initial questionnaire, thereby enhancing the quality of questionnaire development. The selection criteria for correspondence experts included: 1) Over 10 years of experience in nursing management, nursing education, or clinical nursing; 2) a bachelor’s degree or higher; 3) an associate senior or higher technical title; and 4) willingness to participate in the study. Fulfilling two of the first three criteria along with the fourth was sufficient for inclusion. The expert correspondence questionnaire comprised four parts: 1) A letter to experts providing guidance, introducing the study’s background and significance, related concepts, questionnaire collection timeline, and researcher contact information; 2) an expert comment form, including the four-dimensional work performance questionnaire for clinical nurses, instructions for completion, and a scoring system for item importance from ‘very important’ to ‘very unimportant’ on a 5-point scale, with a column for revised opinions; 3) an expert basic information questionnaire capturing age, title, years of experience, and other demographic details; 4) an expert authority questionnaire assessing familiarity with the survey content and the basis for judgment. With consent, experts were asked to evaluate and return the questionnaire via email within two weeks. For indicator screening, items were selected based on the mean importance score and coefficient of variation (CV): Items with a mean importance score > 3.5 and CV ≤ 0.25 were retained [18]. We also considered experts’ opinions and suggestions, and after reviewing the literature and team discussions, decided whether to adjust the items.

2.3. Pre-survey

In September 2024, a pre-survey was conducted using convenience sampling to select 10 clinical nurses from a tertiary hospital in Xinjiang to gather feedback on the clarity of each questionnaire item. All participants, who were registered nurses with at least one year of clinical experience, fully comprehended the content and completed the survey within 15 min. Inclusion criteria for participants included being working registered nurses engaged in clinical practice for at least one year and providing informed consent with a willingness to cooperate. Exclusion criteria encompassed non-clinical staff, those on external training, and intern nurses.

2.4. Assessing questionnaire reliability

2.4.1. Formal investigations

Convenience sampling was employed to select clinical nurses from four tertiary hospitals in the Xinjiang region for the survey conducted from September to November 2024, using the same inclusion criteria as previously described. The sample size required for factor analysis was determined to be 5–10 times the number of questionnaire items [19]. With an initial questionnaire comprising 50 items, the

calculated sample size ranged from 275 to 550, accounting for 10% potential invalid responses. Ultimately, 584 clinical nurses were surveyed, and 549 cases were included for the questionnaire's reliability analysis. The initial questionnaire was uploaded to Questionnaire Star, and data collection was conducted online. Prior to the survey, approval from the hospital nursing departments was obtained, and nurses were briefed on the requirements and precautions for completing the questionnaire. They were instructed to complete all questions in one submission. Questionnaires completed in less than 300 s, with excessively uniform option selection, or exhibiting illogical patterns were excluded, resulting in the elimination of 35 invalid responses and a final inclusion of 549 valid questionnaires, yielding an effective response rate of 94%.

2.4.2. Reliability tests

(1) Internal consistency reliability: Cronbach's alpha coefficient was computed for the total scale and each subscale. Generally, a Cronbach's alpha value greater than 0.9 is considered to indicate high reliability; (2) split-half reliability: The data were divided into two equal parts, and the Spearman-Brown prophecy formula was applied to estimate the reliability of the two halves; (3) test-retest reliability: To assess the scale's stability over time, 20 randomly selected participants were retested after a three-week interval.

2.4.3. Validity tests

(1) Content validity: Expert evaluations of the relevance of questionnaire items to their respective dimensions were conducted. The relevance scores were aggregated to calculate the item-level content validity index (I-CVI) and the scale-level content validity index (S-CVI). Concurrently, the inter-rater agreement (IR) was assessed. Content validity is considered adequate when I-CVI is ≥ 0.78, S-CVI is ≥ 0.80, and IR is ≥ 0.70 [20]; (2) structural validity: Confirmatory factor analysis was employed to evaluate the congruence between the empirical data and the theoretical model. Commonly utilized fit indices include $X^2/df < 3$, RMSEA < 0.08, CFI > 0.9, and SRMR < 0.08 [21].

2.5. Statistical methods

Data were processed and analyzed using SPSS version 27.0 and Amos version 24.0. Continuous measurement data were expressed as the mean ($X \pm SD$), while count data were presented as frequencies. The credibility coefficient and the questionnaire response rate were calculated to assess the reliability and responsiveness of the expert panel. The coefficient of variation and Kendall's W were computed to evaluate the consensus among expert opinions. Item analysis was conducted to screen items based on the following criteria: 1) Critical ratio method: High and low subgroups (defined as the top 27% and bottom 27% of total questionnaire scores, respectively) were compared using independent samples t-tests, and items with non-significant score differences were eliminated; 2) correlation coefficient method: Pearson's correlation coefficients between item scores and total scale scores were calculated, with items considered for removal if the correlation was below 0.4; 3) Cronbach's alpha method: Items were removed if the overall Cronbach's alpha of the scale increased significantly post-removal. The structural

validity of the questionnaire was assessed through exploratory and confirmatory factor analyses. The content validity index was utilized to evaluate the content validity of the questionnaire, with statistical significance set at $p < 0.05$.

3. Results

3.1. Demographic information of consulted experts

A total of 18 experts from regions including Xinjiang, Shanghai, Jiangsu, Guangdong, Gansu, and Hebei were consulted for this study, covering three key areas: nursing management, clinical nursing, and nursing education. The experts' educational backgrounds were as follows: 5 had undergraduate degrees (29.4%), 9 had master's degrees (50%), and 3 had doctoral degrees (16.7%). In terms of professional titles, 2 were intermediate (11.1%), 7 were associate (38.9%), and 8 were full senior (44.4%). Their work experience was distributed as follows: 5 experts had 10–20 years of experience (29.4%), 11 had 20–30 years of experience (64.7%), and 1 had over 30 years of experience (5.9%). For further details, see **Table 1**.

Table 1. Basic information on experts.

sports event	**categorization**	**First round (*n* = 17)**		**second round (of match, or election)**	
		number of people	**Percentage (%)**	**quorum**	**percentage**
distinguishing between the sexes	male	0	0.0	0	0.0
	daughter	17	100.0	17	100.0
education attainment	undergraduate (adjective)	5	29.4	5	29.4
	Master's degree student	9	52.9	9	52.9
	PhD student	321	17.6	3	17.6
(a person's) age	35–45 years	9	52.9	9	52.9
	45–55 years	8	47.0	8	47.0
title	middle level (in a hierarchy)	2	11.8	2	11.8
	deputy high ranking	7	41.1	7	41.1
	high ranking	8	47.0	8	47.0
duties	head nurse	7	41.1	7	41.1
	(Deputy) Director of Nursing	6	35.2	6	35.2
	(Vice) President	2	11.8	2	11.8
	Director of Teaching and Research	1	5.9	1	5.9
	Secretary of the General Party Branch	1	5.9	1	5.9
years of experience	10–20 years	5	29.4	5	29.4
	20–30 years	11	64.7	11	64.7
	> 30 years	1	5.9	1	5.9

3.2. Expert consultation results

In this study, a total of 18 questionnaires were distributed in the first round, with 17 returned, yielding a valid response rate of 94.4%. The positive response

coefficient of the experts was also 94.4%. In the second round, all 17 distributed questionnaires were recovered, achieving a perfect response rate of 100%. The credibility ratio (Cr) was calculated as (0.97 + 0.82)/2 = 0.895, indicating a high level of expert authority and thus enhancing the credibility of the results. After two rounds of expert consultation, the Kendall's W coefficient of concordance for the questionnaire increased significantly from 0.143 to 0.310 (X^2 = 342.092, $p < 0.01$), suggesting an increased level of consensus among the experts. Regarding the concentration of expert opinions, items with a mean importance score greater than 3.5 and a coefficient of variation less than 0.25 were retained. Any items not meeting these criteria were adjusted in consultation with professional opinions and through discussions within the research group.

Results of the First Round of Expert Consultation: Based on the established screening criteria and incorporating the modifications and suggestions provided by the experts, the initial questionnaire underwent the following revisions after deliberation within the research group. A total of 5 items were removed, including "putting patients' interests first, "conscientiously completing work assigned by leaders at all levels", and "caring about the development of the hospital and departments". Additionally, 9 new items were introduced, such as "timely assessment to grasp the dynamic changes in the patient's condition" and "failure to fully respect the patient, infringing upon the patient's physical, informational, social, and other privacy rights." Five items were amended. For example, "endeavoring to help patients solve the difficulties in their lives" was revised to "providing patients with conscientious life care services", and the responsibility for continuity of care services was shifted from the dimension of relational performance to task performance. Furthermore, several expressions were fine-tuned in response to the experts' feedback on select items.

Results of the second round of expert consultation: The second round of expert consultation demonstrated greater consistency compared to the first round, with fewer modifications and suggestions. Two items were deleted: "Able to complete nursing room visits" and "Disobedience to the human resources deployment system". Several items were modified, including reordering the items in Dimension A to align with the definition of nursing services for task performance within that dimension. The item "Able to organize and complete business studies in the department" was revised for clarity, and "Able to organize and participate in the completion of the department's operational learning work" was refined to emphasize the organizational aspect. Additionally, "Possesses critical thinking and is able to identify nursing problems" was adjusted to "Possesses critical thinking and is able to identify nursing problems in a timely manner". Following these revisions, the final "Clinical Nurses' Four-Dimensional Work Performance Questionnaire" comprised four dimensions and 50 items.

3.3. Questionnaire reliability results

3.3.1. Respondent demographics

In this study, a total of 549 clinical nurses were surveyed. The respondents' ages ranged from 22 to 65 years, with a mean age of 35.73 (± 7.56). The majority of

the respondents were female (537, 97.81%), while 12 (2.19%) were male. Regarding educational background, 8 (1.46%) had a secondary school education, 84 (15.30%) had an associate degree, 450 (81.97%) held a bachelor's degree, and 7 (1.28%) had a master's degree. In terms of job titles, the distribution was as follows: 61 (11.11%) staff nurses, 202 (36.79%) nurse practitioners, 246 (44.81%) charge nurse practitioners, 22 (4.00%) associate nurse practitioners, and 17 (3.10%) chief nurse practitioners. In terms of work experience, 70 (12.75%) had worked for 5 years or less, 145 (26.41%) for 5–10 years, 128 (23.31%) for 10–15 years, and 206 (37.52%) for more than 15 years. Regarding job levels, the distribution across the six tiers from N0 to N5 was as follows: 11 (2.00%) at N0, 73 (13.30%) at N1, 200 (36.43%) at N2, 219 (39.90%) at N3, 32 (5.83%) at N4, and 14 (2.55%) at N5.

3.3.2. Project analysis

Critical Ratio Method: In this study, the top 27% of nurse work performance questionnaire scores were designated as the high group, and the bottom 27% as the low group. Independent samples t-tests were conducted to assess the critical ratio (CR) of each item. Based on the analysis, the items "Take the initiative to care for patients from both physiological and psychological aspects to meet their needs", "Be concerned about the development of hospitals and departments and able to make timely suggestions for their development", and "The occurrence of nurse-patient conflicts resulting in damage to patients' rights and interests" had CR values of 2.61, 0.85, and 0.09, respectively, with p values greater than 0.05. These items did not demonstrate good discriminative power and were therefore excluded.

3.3.3. Reliability analysis

The overall Cronbach's alpha coefficient for the questionnaire was 0.962, indicating high internal consistency. The split-half reliability coefficient was 0.795, and the test-retest reliability coefficient was 0.964, both of which are within acceptable ranges for reliability. As detailed in **Table 2**, the Cronbach's alpha coefficients for all dimensions were also at a satisfactory level.

Table 2. Overall and dimensional reliability of the questionnaire.

sports event	reliability	retest reliability	half confidence
General questionnaire	0.962	0.964	0.734
Mission performance	0.967	0.904	0.968
Relationship performance	0.899	0.920	0.895
Adaptive performance	0.953	0.931	0.953
Anti-production performance	0.909	0.964	0.906

3.3.4. Comprehensive validity testing

Content validity

Based on the experts' assessments of the relevance of each questionnaire item, the item-level content validity index (I-CVI) ranged from 0.82 to 1. The total content validity index for the questionnaire (S-CVI/UA) was 0.71, and the average content validity index (S-CVI/Ave) was 0.97. These values indicate that the questionnaire possesses good content validity.

Structural validity

(1) Exploratory factor analysis: Bartlett's test of sphericity for the scale indicated that the data were suitable for factor analysis ($p < 0.001$), and the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was 0.965, further supporting the factorability of the correlation matrix. Through principal component analysis, factors with eigenvalues greater than 1 were extracted, yielding 4 factors that accounted for 63.768% of the cumulative variance. After orthogonal rotation using the maximum variance method, the loadings of each item on the corresponding common factor were all greater than 0.4. The structure of each item and dimension of the questionnaire corresponded to the initial design, with specific details presented in **Table 3**.

(2) Validation factor analysis: The maximum likelihood method was employed to analyze and fit the model. The questionnaire's four dimensions—task performance, relationship performance, adaptive performance, and counterproductive performance—were subjected to confirmatory factor analysis. All item factor loadings were greater than 0.6, suggesting strong relationships between the items and their respective factors. Additionally, the good model fit indices indicated that the theoretical design of the questionnaire closely aligned with the empirical data. For detailed results, refer to **Table 4.**

(3) The results of the convergent and discriminant validity tests demonstrate that the questionnaire possesses strong convergent validity, with the average variance extracted (AVE) for the four dimensions exceeding 0.5 and the composite reliability (CR) surpassing 0.9. Additionally, the square root of the AVE for each dimension is greater than the absolute value of the Pearson correlation coefficients with other dimensions, indicating that the questionnaire also exhibits strong discriminant validity. For detailed results, refer to **Table 3**.

Table 3. Parameters for factor analysis of the questionnaire.

dimension (math.)	Content of entries	pearson (r)	Adaptive performance	Mission performance	Anti-production performance	Relationship performance	AVE	CR
A- take up a post affair spin (hemp etc) efficacy	A2. Inquire about patients' needs and questions and resolve or answer them in a timely manner	0.891**	0.287	**0.896**	0.216	0.199	0.682	0.969
	A3. Timely assessment to keep track of the patient's dynamic changes in condition.	0.733**	0.244	**0.75**	0.159	0.14		
	A4. Follows medical orders to the letter and provides accurate and timely patient care.	0.711**	0.229	**0.744**	0.146	0.13		
	A5. Closely observe changes in the patient's condition and address any abnormalities promptly.	0.734**	0.204	**0.757**	0.181	0.181		
	A6. Do what you can to help patients relieve their pain.	0.726**	0.256	**0.74**	0.165	0.11		
	A7. Take the initiative to care for patients physically and psychologically and meet their needs	0.753**	0.243	**0.766**	0.148	0.176		
	A8. provides service to patients in a cordial manner and establishes a good nurse-patient relationship.	0.722**	0.224	**0.736**	0.219	0.117		
	A9. carefully provide personalized health education to patients and families.	0.716**	0.193	**0.769**	0.123	0.185		
	A10.Carefully complete continuity of care services.	0.738**	0.221	**0.775**	0.158	0.128		
	A11. Proficiency in basic nursing knowledge and nursing practice.	0.710**	0.192	**0.734**	0.19	0.185		
	A12. Proficiency in specialty nursing knowledge and nursing practice.	0.729**	0.241	**0.742**	0.144	0.173		
	A13. Planned care is organized and efficiently completed.	0.729**	0.225	**0.748**	0.163	0.161		
	A14. Have the ability to provide evidence-based care and make sound decisions based on the best evidence to bring better care to patients.	0.727**	0.204	**0.757**	0.213	0.139		
	A15. Have good nurse-patient communication skills.	0.725**	0.227	**0.758**	0.164	0.137		
	A16. Ability to conduct patient safety/risk assessments to safeguard patients.	0.734**	0.25	**0.737**	0.154	0.157		

Table 3. (*Continued*).

dimension (math.)	Content of entries	pearson (r)	Adaptive performance	Mission performance	Anti-production performance	Relationship performance	AVE	CR
B-Relationship performance	B1. Learn from each other and make progress with colleagues.	0.548**	0.116	0.291	0.037	**0.899**	0.633	0.909
	B2. works in solidarity with coworkers in the delivery of nursing care.	0.451**	0.092	0.244	0.036	**0.733**		
	B3. Work with coworkers to maintain medical order and create a favorable environment for patient care.	0.442**	0.129	0.173	0.054	**0.75**		
	B4. be able to perform well in teaching nursing students and nurses.	0.431**	0.082	0.164	0.087	**0.793**		
	B5. be able to organize and participate in the completion of the section's business studies.	0.449**	0.098	0.204	0.066	**0.771**		
	B6. conscientiously participate in hospital quality control.	0.417**	0.094	0.217	-0.02	**0.719**		
C-Adaptive performance	C1. Actively participate in various trainings organized by the hospital and department.	0.780**	**0.933**	0.272	0.08	0.082	0.562	0.955
	C2. takes initiative in learning new nursing theories and techniques.	0.564**	**0.684**	0.167	0.11	0.083		
	C3. ability to apply new nursing theories and techniques to clinical work.	0.549**	**0.732**	0.136	0.04	0.027		
	C4. Insist on continuing education and learning, and grasp the opportunity to upgrade their qualifications.	0.584**	**0.709**	0.213	0.067	0.005		
	C5. possesses critical thinking and is able to identify nursing issues.	0.580**	**0.716**	0.216	0.021	0.039		
	C6. Be able to propose creative solutions to nursing problems.	0.585**	**0.735**	0.153	0.097	0.074		
	C7. has a willingness to do research and to take on innovative tasks.	0.590**	**0.702**	0.177	0.094	0.096		
	C9. Is able to handle other emergencies.	0.565**	**0.714**	0.186	0.021	0.045		
	C10. Be able to control his/her emotions in the face of pressure from colleagues, patients and patients' families.	0.565**	**0.680**	0.166	0.034	0.124		
	C11. Identifies with the values, spirit, and philosophy of the hospital and department.	0.563**	**0.749**	0.15	0.022	0.078		
	C12. actively participates in activities related to the promotion and building of hospital and departmental culture.	0.567**	**0.703**	0.198	0.036	0.031		

Table 3. (*Continued*).

dimension (math.)	Content of entries	pearson (r)	Adaptive performance	Mission performance	Anti-production performance	Relationship performance	AVE	CR
C-Adaptive performance	C13. Be able to guide his/her behavior by the values and philosophy of the hospital and department.	0.594**	**0.681**	0.224	0.029	0.12		
	C14. Strictly implement all operating procedures such as the checking system.	0.589**	**0.744**	0.153	0.05	0.108		
	C15. Continuous Improvement of Nursing Care.	0.607**	**0.740**	0.183	0.071	0.086		
	C16. Ability to apply the nursing process to provide holistic patient care.	0.613**	**0.688**	0.218	0.11	0.083		
	C17. Is actively involved in public service such as popularization of science and volunteer clinics.	0.577**	**0.726**	0.176	0.067	0.028		
	C18. Closely observe the patient's condition, and promptly resuscitate in case of accidents.	0.552**	**0.719**	0.131	0.093	0.048		
D-Anti-production performance	D1. There were late arrivals, early departures and absences.	0.582**	0.12	0.302	**0.902**	0.037	0.624	0.917
	D2. Disobedience to organizational working arrangements.	0.424**	0.093	0.174	**0.738**	0.012		
	D3. Complains a lot and is negative.	0.427**	0.056	0.174	**0.783**	0.053		
	D4. Poor communication with patients and families and failure to adequately safeguard the patient's right to information.	0.410**	0.023	0.221	**0.733**	0.029		
	D6. Poor solidarity and collaboration with colleagues.	0.409**	0.07	0.156	**0.774**	0.04		
	D7. Inappropriate grooming or speech.	0.468**	0.128	0.205	**0.741**	0.046		
	D8. Failure to fully respect the patient and invasion of the patient's physical, informational, and social privacy.	0.464**	0.084	0.242	**0.716**	0.063		

footnote. **: The *P* value was less than 0.01.

Table 4. Questionnaire validation factor model fit indices table.

statistic	standard of judgment	Model fit index
X^2/df	< 3	1.096
GFI	> 0.9	0.95
RMSEA	< 0.10	0.013
RMR	< 0.05	0.024
CFI	> 0.9	0.995
NFI	> 0.9	0.95

4. Discussion

4.1. Rationale for developing a four-dimensional scale for assessing nursing job performance

Nurses' work performance encompasses both the process and outcomes of their duties, serving as a crucial basis for the development of nursing human resource management [22]. It plays a significant role in the evolution of nursing careers. The National Nursing Career Development Plan advocates for the implementation of scientific performance appraisal tools to fully engage nurses' motivation. Previous research on nurses' work performance has primarily utilized scales such as the three-dimensional performance workload scale developed by Ma Yanlan and Yang Xiaoli [23], and the six-dimensional scale of nursing work performance by Dr. Schwirian. However, these tools were developed earlier and do not reflect the latest performance management theories or the current realities of nursing work. This study introduces a four-dimensional work performance questionnaire for clinical nurses, which includes Adaptive Performance in addition to Task Performance, Relationship Performance, and Counterproductive Performance [24]. Adaptive Performance refers to an individual's capacity to adapt and cope with new situations in a dynamic work environment. In the contemporary era, the adaptive performance of healthcare workers is particularly critical. Nurses with higher levels of adaptive performance can more rapidly acquire new technologies, enhance nursing service efficiency, and adjust their behavioral and functional work styles flexibly to improve the quality of nursing services. Enhancing adaptive performance can assist hospitals in more rational human resource allocation. Investigating nurses' work performance across four dimensions provides a comprehensive understanding of their work, enabling the development of targeted interventions to ultimately enhance work performance and improve the quality of nursing services.

4.2. Scientific and reliable questionnaire development

This study is grounded in Koopmans' four-dimensional performance structure theory and builds upon a comprehensive review of relevant literature on the performance appraisal index system for nurses in China [8–15]. The questionnaire was compiled by consulting with nursing management experts, ensuring a rational structural layout, clear logical sequence, and significant clinical application value. Through literature research, Delphi expert correspondence, and research group discussions, a final set of 45 nurse performance questionnaire items was developed.

The experts consulted through the Delphi method were selected from six provinces and cities across China, all holding intermediate or higher titles and possessing over 10 years of experience in clinical nursing, nursing management, and nursing education. They served as department head nurses, directors of nursing departments, or university professors, providing insights from both clinical nursing and nursing management perspectives. In the two rounds of correspondence questionnaires, the experts' credibility coefficients were 0.90 and 0.92, respectively, with positive response coefficients of 94.4% and 100%. The Kendall's coefficient of concordance increased from 0.143 to 0.310 across the two rounds, and the coefficient of variation for the questionnaire dimensions and each item was less than 0.25. In summary, the experts involved in this study were representatively distributed across multiple provinces and cities, demonstrating good positivity, authority, and coordination. The results of the Delphi correspondence were highly credible and reliable, aligning with established norms and exhibiting strong scientific validity and reliability.

4.3. The questionnaire has good reliability and validity

In this study, the initial version of the questionnaire was administered to 549 nurses in tertiary hospitals in Xinjiang to assess its reliability. A series of analyses, including the critical ratio method, correlation coefficient method, factor analysis method, and internal consistency reliability analysis, were conducted on the initial questionnaire following expert consultation. Item analysis resulted in the retention of 45 items, and the reliability of the four-dimensional job performance questionnaire for clinical nurses was examined using various methods. The results indicated that the overall internal consistency coefficient of the questionnaire was 0.962, with dimension-specific coefficients ranging from 0.899 to 0.967, and the re-test reliability of the total questionnaire was 0.964. The exploratory factor analysis demonstrated that all item factor loadings exceeded 0.4. Furthermore, the confirmatory factor analysis showed that all model fit indices met the criteria, and internal consistency was satisfactory. The overall content validity of the questionnaire was $\geq 90\%$. During the reliability and validity analysis, we paid special attention to the performance of items involving biological knowledge and skills. For example, some biology-related items such as "understanding the impact of genetic test results on patient care plans" and "correct interpretation of biological test results (such as blood glucose, blood lipids)" performed poorly in the reliability analysis. This may be because there are differences in the understanding and application of these biological concepts in nursing practice, or because the questionnaire failed to accurately cover the complexity of biological content. In order to further improve the reliability and validity of these items, we recommend that in future studies: strengthen training and education to ensure that nurses have sufficient basic biological knowledge and skills training, especially in the application of new technologies such as genetic testing and immunotherapy; refine questionnaire items, and for complex biological concepts, consider splitting questions into more specific sub-items to ensure that each item can clearly reflect the actual ability of nurses; multi-center validation, expand the scope of research, and validate across multiple regions and hospitals to better capture the application of biological knowledge in

different practice environments. In addition, biology-related items in counterproductive performance, such as "non-compliance with aseptic operation principles" and "behavior that ignores patients' biological needs", showed high discrimination in the reliability and validity analysis. This suggests that these items are effective in identifying behaviors that are not consistent with biological nursing principles, emphasizing the importance of nurses following biological principles in their daily work. In terms of validity, we evaluated whether the questionnaire truly measured the biological dimensions related to nursing performance from a biological perspective. In particular, in the confirmatory factor analysis, we noted the fit of the factors related to biology. For example, items such as "understanding the impact of genetic test results on patient care plans" and "correctness of interpreting biological test results (such as blood glucose and blood lipids)" in task performance, and "ability to cope with the development of biomedical technology" in adaptive performance. If the fit index of these biological-related factors is not ideal, we will further analyze the reasons and consider the following two aspects: First, the mismatch between biological theory and questionnaire construction may be due to the fact that the questionnaire design fails to fully cover the latest biological theories or biological concepts involved in actual nursing work. To improve this problem, future studies can combine the latest biological research results to re-examine and adjust the questionnaire items to ensure that they are consistent with current biological knowledge and technological development. Second, the performance of biological factors in actual nursing work is different from expectations. In actual nursing work, nurses may face different challenges and limitations, resulting in the performance of some biological factors being different from expectations. For example, the application of new technologies may have problems such as insufficient training or limited resources, which affects the actual operation effect of nurses. Therefore, it is recommended that in future validation studies, attention should be paid not only to the theoretical fit of the questionnaire, but also to the actual performance of biological factors and their influencing factors in combination with the biological application in the actual nursing environment. In conclusion, the questionnaire developed in this study exhibits good reliability and validity, with items that possess strong representativeness, differentiation, and discriminative abilities.

5. Conclusions

The Delphi expert correspondence method inherently carries a degree of subjectivity, and it is not feasible to incorporate all expert opinions into the final selection. Additionally, due to temporal and geographical constraints of this study, the reliability test was conducted exclusively with clinical nurses from a tertiary hospital in Urumqi City. Future studies should aim to validate the questionnaire across multiple centers and regions to enhance the generalizability and content validity of the instrument.

Author contributions: Conceptualization, PL and HC; methodology, HC; software, XZ; validation, HC, TZ and XZ; formal analysis, HC; investigation, HC; resources,

MH; data curation, TZ; writing—original draft preparation, HC; writing—review and editing, HC; visualization, HC; supervision, PL; project administration, MH; funding acquisition, PL. All authors have read and agreed to the published version of the manuscript.

Ethical approval: This study was conducted in accordance with the Declaration of Helsinki and was approved and ratified by the Ethics Committee of the Medical School and the First Affiliated Hospital of Shihezi University (Grant number: KJ2024-008-01). Informed consent was obtained from all participating subjects before the study was conducted.

Conflict of interest: The authors declare no conflict of interest.

References

1. Su B, Du D, Jia J, et al. Research on the current situation of nursing human resources and their allocation in China. China Health Policy Research. 2018; 11(12): 56–61.
2. Li X, Luo M, Gao X, et al. Application and thinking of nursing performance management system. Chinese Journal of Modern Nursing. 2015; 21(1): 75–77.
3. China Nursing Management. National Nursing Career Development Plan (2021–2025). China Nursing Management. 2022; 22(6): 801–804.
4. General Office of the National Health and Health Commission. Circular of the General Office of the National Health and Wellness Commission on Further Strengthening Nursing Care in Medical Institutions. Bulletin of the National Health Commission of the People's Republic of China. 2020; 8: 10–13.
5. Ma Y, Wang J, Wang H, et al. Development of a performance evaluation scale for hospital nurses. Journal of Nursing Management. 2007; 3: 14–16.
6. Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. Journal of Advanced Nursing. 2000; 32(4): 1008–1015.
7. Calvio Suárez C, Martinez-Rodriguez AL, Moreira-Alvarez D, et al. P021 Higher TNFα and IL-6 mucosal levels in ulcerative colitis patients with more severe endoscopic activity. Journal of Crohn's and Colitis. 2024; 18(1).
8. Davaanyam D, Seol SI, Oh SA, et al. Hepatocyte activation and liver injury following cerebral ischemia promote HMGB1-mediated hepcidin upregulation in hepatocytes and regulation of systemic iron levels. Experimental & Molecular Medicine. 2024; 56(10): 2171–2183.
9. Koopmans L, Bernaards CM, Hildebrandt VH, et al. Conceptual Frameworks of Individual Work Performance: A Systematic Review. Journal of Occupational & Environmental Medicine. 2011; 53(8): 856–866.
10. Zhang X, Zheng J, Xu H, et al. Construction of evaluation index system for nursing individual performance appraisal in a tertiary general hospital. China Medical Management Science. 2019; 9(3): 44–49.
11. Bao H, Wang X, Ben H, et al. Research on the construction of nursing performance evaluation index system in public hospitals—Based on hierarchical analysis method. Health Soft Science. 2024; 38(2): 36–40.
12. Fang G, Liu Y, Gan Z, et al. Application of a tiered performance appraisal model for nursing units based on RBRVS and DRGs. Journal of Nursing Management. 2023; 23(3): 210–214.
13. Liu Y, Zhang P, Yu X, et al. Developmental performance evaluation model and index construction for clinical nurses based on ERG theory. Journal of Nursing. 2023; 38(11): 58–62.
14. Wei Y, Xi Z, Tian S, et al. A focus group interview study on the performance appraisal index of operating room nurses in a specialized oncology hospital based on the four-dimensional model of individual performance. Journal of Nursing. 2024; 31(3): 30–35.
15. Zhang R, Zhang J, Ma J, et al. Research on the construction of performance appraisal index system for clinical health care personnel. China Geriatric Healthcare Medicine. 2023; 21(2): 65–67.
16. Yang X, Zhu Y, Liu Y, et al. Construction of performance assessment indexes for outpatient triage nurses. Journal of Nursing. 2019; 34(15): 62–65.
17. Zeng Z, Zhou S, Liu M. Research progress on assessment tools related to occupational fatigue in nurses: A traditional review.

Frontiers in Public Health. 2024; 12: 1508071.
18. Zhao Y, He L, Hu J, et al. Using the Delphi method to establish pediatric emergency triage criteria in a grade a tertiary women's and children's hospital in China. BMC Health Services Research. 2022; 22(1): 1154.
19. MacCallum RC, Widaman KF, Preacher KJ, Hong S. Sample size in factor analysis: the role of model error. Multivariate Behavioral Research. 2001; 36(4): 611–637.
20. Shi J, Mo X, Sun Z. Application of content validity index in scale development. Journal of Central South University (Medical Edition). 2012; 37(2): 49–52.
21. Zhang Y, Yang R, Peng X, et al. Principles and models of validated factor analysis and its application in the evaluation of scale structural validity. Chinese Journal of Health Management. 2022; 16(6): 428–432.
22. Yang XM, Chen YM. Strengthening nursing staff work performance management Enhancing nursing human resource management level. Journal of Bengbu Medical College. 2017; 42(1): 1–2.
23. Yang X, Su X, Zhang J, et al. Evaluation of nurses' job performance and its influencing factors. Nursing Research. 2006; 7: 631–633.
24. Pulakos ED, Arad S, Donovan MA, et al. Adaptability in the workplace: development of a taxonomy of adaptive performance. The Journal of Applied Psychology. 2000; 85(4): 612–624.