BMC Nursing

RESEARCH

Open Access

Measuring honesty in nursing: scale development and validation

Mohammad Mehdi Mohammadi[1], Amirhossein Naghibzadeh[2], Hadis Mosafer[3] and Mohammad Mohammadi[2*]

Abstract

Background Honesty is considered a desirable trait that can impact happiness and job performance. The clinical environment can present ethical challenges for nurses. A lack of honesty can lead to significant issues for both nurses and patients. Existing tools for measuring honesty have limitations and do not focus specifically on nurses. This study aimed to design and psychometrically evaluate a scale to measure honesty among nurses.

Methods This methodological study involved the design and validation of an honesty measurement tool for nurses. The study employed an inductive-deductive approach to generate items. The face (10 nurses), content (10 nursing experts), and construct validity (320 hospital nurses) of the scale were assessed, along with reliability using internal consistency and stability methods. The Item Impact Method was used for quantitative face validity determination. Content validity was assessed using qualitative and quantitative methods, with the participation of researchers and nursing experts. Cronbach's alpha was used to assess internal consistency, and the test-retest method and Intraclass Correlation Coefficient (ICC) were used to assess reliability. Exploratory factor analysis (EFA) was used to assess construct validity.

Results The study resulted in an 8-item tool with two dimensions: behavioral honesty and verbal honesty. The tool explained 51.034% of the total variance. Cronbach's alpha for the entire scale was 0.823, and McDonald's omega was 0.898. The Intraclass Correlation Coefficient (ICC) was reported as 0.918, with a 95% confidence interval of 0.827 to 0.961.

Conclusion The developed tool is a reliable and valid instrument for measuring honesty among nurses. This study contributes to the field by providing a tool to measure honesty in nurses. The tool's scoring system effectively distinguishes between different levels of honesty in nursing practice, with elevated scores reflecting greater adherence to honest professional conduct. This tool can be used by managers and organizational decision-makers to enhance the quality of management planning.

Clinical trial number Not applicable.

Keywords Professional integrity, Honesty, Nursing, Psychometrics

*Correspondence:
Mohammad Mohammadi
m456.mohammadi@gmail.com
[1]Department of Medical-Surgical Nursing, School of Nursing & Midwifery, Kermanshah University of Medical Sciences, Kermanshah, Iran
[2]Student Research Committee, Kermanshah University of Medical Sciences, Kermanshah, Iran
[3]Department of Geriatric and Psychiatric Nursing, School of Nursing & Midwifery, Kermanshah University of Medical Sciences, Kermanshah, Iran

Background

Honesty is widely regarded as a valued ethical virtue [1]. Studies have shown that honesty is associated with psychological well-being and life satisfaction [2]. Honest individuals often experience a greater sense of authenticity and inner coherence, which can contribute to feelings of happiness and contentment [3]. Eijkholt (2018) defines honesty as the extent to which individuals report they could be honest in their interactions [4]. Honesty is upheld as a cornerstone principle in numerous cultures and societies globally [5].

Promoting honesty in the workplace yields significant positive outcomes. When individuals consistently act with integrity, it fosters a culture of trust, which in turn enhances collaboration and productivity [6, 7]. Honest communication reduces misunderstandings and conflicts, streamlining operations and boosting morale. Furthermore, a reputation for honesty attracts and retains top talent, as employees are drawn to ethical work environments. Ultimately, workplace honesty strengthens the organization's credibility with clients and stakeholders, leading to improved relationships and long-term success [6–9].

As a cornerstone of human values, honesty demands paramount attention from all members of society, particularly those within the medical sciences, including nursing. These professions bear the weighty responsibility of directly impacting the health and lives of individuals. However, the clinical environment often presents ethical dilemmas and challenges that can strain the principles of honesty and ethical conduct for nurses. A significant disparity between expected and actual clinical realities increases the likelihood of errors and potential challenges [10, 11].

Nursing, a profession rooted in trust and teamwork, necessitates the presence of honesty as a foundational element for effective healthcare delivery. However, there currently exists a significant gap in the tools available to measure this ethical component specifically among nurses [12]. Developing and validating a reliable instrument to evaluate nurses' honesty is crucial to address this gap. Such a tool will not only provide essential data for understanding honesty within nursing but will also offer actionable insights to enhance management strategies aimed at promoting integrity in clinical settings. By addressing these needs, this research can ultimately support better planning, decision-making, and team dynamics within healthcare environments [12, 13].

In a search for existing tools and questionnaires worldwide related to the concept of honesty in nursing, we come across the study by Muramatsu et al. which focused on the design and psychometric evaluation of the Ethical Sensitivity Questionnaire among nursing students. In this study, a scale was developed to measure the ethical sensitivity of students enrolled in 10 Japanese universities. Notably, this study concentrated on students rather than practicing nurses, and the concept under investigation was ethical sensitivity, which differs from the concept of honesty in nursing [14].

In a study conducted by Wote et al., the role of honesty among elementary school principals was examined and interpreted. This study focused on the impact of principals' honesty on teachers and the subsequent influence on teaching practices. The data revealed a positive and direct correlation between honesty and the performance of both teachers and students. While this study clearly demonstrated the positive effects of honesty, it is important to note that it had no direct relevance to the nursing field. The concept of honesty was explored within the context of education, using a sample population of school principals rather than nurses. Consequently, this study cannot be considered a reliable source for addressing the challenges and concepts specific to the nursing profession [15].

In another study, conducted across Iran, Nigeria, and New Zealand, and titled the "Cross-Cultural Academic Integrity Questionnaire" (CCAIQ), the focus was on three concepts: "compliance, collusion, and cheating." This study examined academic honesty in higher education settings, excluding nursing programs. In other words, the scale was not designed to measure honesty within the nursing profession. Instead, it was used to assess a sample of higher education students, which cannot provide an accurate and reliable basis for measuring the level of honesty among nurses [16].

Despite extensive efforts to identify existing measurement tools for honesty, a significant gap remains in addressing this concept specifically within the nursing profession. Prior studies either focus on related but distinct concepts or target different populations, emphasizing the urgent need for a specialized instrument tailored for nurses. This research seeks to bridge this gap by designing and psychometrically validating a scale for assessing honesty among nurses in Iran. The outcome of this study is expected to provide a robust tool for organizational leaders and healthcare managers to make evidence-based decisions. By enabling precise assessments of nurses' honesty, this tool will contribute to improved management planning and strategies, fostering an ethical and trustworthy healthcare environment that benefits personnel, patients, and the healthcare system as a whole.

Method

Study design and item generation

This methodological study aimed to design and validate a scale to measure honesty among nurses. The study was conducted from November to December 2023. The development of items for the scale followed a systematic

approach. An inductive-deductive methodology was employed [17], incorporating insights from qualitative interviews with practicing nurses and a review of relevant literature. Ten nurses with diverse professional experiences participated in the item generation process, ensuring the initial item pool reflected practical realities and professional ethics in nursing. To ensure robust psychometric evaluation, the study adhered to established guidelines, including the COSMIN (COnsensus-based Standards for the selection of health Measurement INstruments) framework. Psychometric properties, such as face validity, content validity, construct validity (via exploratory factor analysis), and reliability (internal consistency and test-retest reliability), were carefully chosen to validate the tool comprehensively. These properties were selected based on their relevance to the tool's intended purpose, with detailed procedures for each outlined in the methodology. The psychometric properties of the developed scale were evaluated based on Classical Test Theory (CTT) [18].

Face validity

Quantitative determination of face validity

After revisions and necessary changes based on the feedback provided by 10 nurses participating in the study, the Item Impact Method was used as a basis for determining quantitative face validity, following the steps outlined in Table 1.

Qualitative determination of face validity

To conduct this phase of the study, the questionnaire items were evaluated based on three criteria: difficulty level (difficulty in understanding terms and words), irrelevancy (potential irrelevance of items to the dimensions of the scale), and ambiguity (the possibility of misinterpretations or deficiencies in the meanings of words). In this regard, the population for whom the scale was designed (nurses) was selected as participants for this phase, and their responses were used to modify and revise the items.

Content validity

In the present study, both qualitative and quantitative methods were employed to determine content validity [19, 20]:

Qualitative determination of content validity

To assess the qualitative content validity, ten researchers and experts with a background in nursing scale development were invited to review the scale. These individuals were selected from the faculty members with a history of research and studies in the field of nursing. The reviewers were asked to provide feedback based on the following criteria: grammar, appropriate wording, item allocation, scaling, simplicity, and clarity.

Quantitative determination of content validity

The content validity of the scale was quantitatively assessed using two methods: the Content Validity Ratio (CVR) and the Content Validity Index (CVI). First, the CVR was calculated, and based on its results, items were potentially removed. Subsequently, the CVI, utilizing the adjusted kappa statistic (K*) [19–22], was calculated. The CVR calculation followed the steps outlined in Table 2, while the CVI calculation using K* followed the steps in Table 3. In order to calculate K*, the steps in Table 3 were followed.

Initial reliability assessment

Cronbach's alpha was used to assess the internal consistency of the scale. Thirty nurses working in educational hospitals in Kermanshah were selected using a convenience sampling method. To conduct item analysis, the loop method was employed, and the corrected item-total correlation was calculated. Items with a correlation coefficient below 0.3 with the total item score were removed. Additionally, the change in alpha after removing each item was evaluated [21].

Construct validity

To assess construct validity, exploratory factor analysis (EFA) was conducted. The target group (selected nurses) completed the scale, and data was collected using a printed questionnaire. In addition to demographic data, the questionnaire used a 5-point Likert scale (always,

Table 1 Steps of the item impact method for face validity

Step	Description
1	The questionnaire administered to the target population in this study employed a five-point Likert scale (5 = very suitable, 4 = relatively suitable, 3 = unsuitable, 2 = very unsuitable).
2	Based on the Likert scale mentioned above, a score was assigned by ten nurses to each item in the questionnaire. The "item impact score" was calculated separately for each item using the following formula: *Item Impact Score = Frequency (%) x Suitability (Importance)* According to the above formula, Frequency refers to the percentage of individuals who rated the item as 4 or 5, and Suitability represents the average score assigned to that item.

A score of ≥ 1.5 indicated that the item was considered suitable, while a score of < 1.5 suggested that the item required revision and refinement

Table 2 Steps for calculating CVR in the present study

Step	Description
1	A Description of the Instrument's Objectives and Operational Definitions
2	To determine the necessity of the items based on a three-point Likert scale (1. Essential, 2. Useful but not essential, 3. Not necessary), 12 experts (experienced faculty members with a background in nursing and instrument development) were asked to participate.
3	To calculate content validity, the following formula was used: $CVR = \frac{n_e - (N/2)}{N/2}$ In this formula, n_e represents the number of experts who selected the "essential" option, and N represents the total number of experts.
Note: The minimum acceptable numerical value for CVR, as determined by the Lawshe table and calculated with a 95% confidence interval (CI), was 0.56 given a panel of 12 experts. Accordingly, items with a CVR below 0.56 were eliminated. The remaining items were assessed for content validity index (CVI) using the K statistic.*	

often, sometimes, rarely, never) to assess the study population. Recommendations for questionnaire validation suggest a range from as few as five to ten respondents per item to absolute minimums of 200 to 300 participants. Considering these recommendations, a minimum sample size of 200 was selected based on rules of thumb. Recognizing the possibility of excluding some samples based on exclusion criteria, the study population was initially estimated to be 225 [23, 24]. Sampling was conducted using a convenience sampling method, with the inclusion criterion being employment in a hospital and direct patient care. Indifferent respondents and those with a standard deviation of less than 0.2 in their responses, as well as univariate and multivariate outliers, were considered as exclusion criteria. To eliminate these, SPSS version 16 and exploratory factor analysis were used. In the first stage, sample adequacy was assessed using Bartlett's test of sphericity and Kaiser-Meyer-Olkin (KMO) measure. For the KMO measure, a value greater than 0.5 is considered acceptable, but values greater than 0.7 are more desirable. In this study, a minimum acceptable value of 0.7 was considered [21].

This study employed factor analysis to identify underlying factors. To determine the number of factors to retain, eigenvalues and the scree plot were examined. Additionally, parallel analysis was used to make the final factor selection. In this method, only factors with eigenvalues greater than 1 were retained, and the rest were discarded [20, 21]. For each eigenvalue, a 95% confidence interval was calculated. The formula for calculating the confidence interval is as follows [25]:

$$l_i \pm z^* \left(\sqrt{\frac{2l_i^2}{n}} \right)$$

In this context, l_i represents the eigenvalue and n represents the sample size. Given a 95% confidence interval, the value of Z is 1.96. This means that with 95% confidence, the eigenvalue will be greater than or equal to 1.

Table 3 Steps for calculating K* in the present study

Step	Description
1	Initially, the Individual Content Validity Index (I-CVI) for each item was calculated using the conventional method proposed by Waltz and Bausell. To this end, a panel of 12 experts (experienced university professors in instrument development, and nursing) rated each item in terms of its "relevance" on a four-point Likert scale ranging from 1 (not relevant) to 4 (highly relevant). The I-CVI was calculated as the ratio of the number of experts who rated an item as 3 or 4 to the total number of experts. $$\text{CVI} = \frac{\text{number of raters giving a rating of '3' or '4'}}{\text{total number of raters}}$$
2	In the next step, the probability of chance agreement (Pc) was calculated using the following formula: $$P_C = \left[\frac{N!}{A!(N-A)!}\right] \times 0.5^N$$ Where A represents the number of experts who rated the items as 3 or 4, and N represents the total number of experts.
3	In the final stage, the adjusted kappa statistic (K*) was computed using the following formula**:** $$K^* = \frac{I-CVI-Pc}{1-Pc}$$
Note: Polit suggested that the interpretation of K* is as follows: a score between 0.40 and 0.59 is considered poor, 0.60 to 0.74 is considered good, and more than 0.74 is considered excellent; based on this, in this study, items with a score higher than 0.74 were considered relevant and excellent. Moreover, based on the average of I-CVIs, S-CVI/Ave (Scale- Content Validity Index/Average), the content validity of the entire instrument S-CVI (Scale-Content Validity Index) was calculated. A minimum numerical value of 0.90 was considered the acceptance criterion for S-CVI/Ave.	

The scree plot was generated using simulated data and the 95th percentile, using JASP version 0.14. Based on this plot, the extracted factors have observed eigenvalues higher than the 95th percentile of the simulated data.

This study aimed to enhance the quality and interpretability of the factor structure assessment. To achieve this, after extracting the factors, Varimax rotation, an orthogonal rotation method, was employed. The factors were named based on the corresponding items [26]. To determine the minimum factor loading required for an item to remain in a factor, the following formula was used with a 99% confidence interval: CV = 5.152 ÷ √(n-2), where n is the sample size [27]. The minimum factor

loading considered for inclusion in the study was approximately 0.3.

Identifying indifferent respondents

This study used the standard deviation of each participant's responses to identify and exclude samples with indifferent responses. Responses with a standard deviation less than 0.2 were removed from the study. To perform this task, Excel 2007 and a custom function based on standard deviation were utilized.

Examining normal distribution and univariate and multivariate outliers

In this study, skewness (±3) and kurtosis (±7) were used to examine the normality of the data. Box plots were used to identify univariate outliers. Samples with outliers were more likely to be excluded from the study. To find multivariate outliers, the Mahalanobis d-square test was used with Amos version 25. Samples with $P < 0.001$ were more likely to be excluded from the study.

Final reliability

In this study, internal consistency was assessed using Cronbach's alpha, and stability was evaluated using the test-retest method. This means that, in the first stage, Cronbach's alpha was calculated, and in the second stage, temporal stability was examined.

Cronbach's alph

In this study, Cronbach's alpha and Average Inter-item Correlation (AIC) were used to assess internal consistency, and McDonald's omega was used to assess stability. Cronbach's alpha and AIC were calculated with the participation of 30 nurses. McDonald's omega was calculated using the results of exploratory factor analysis. In this study, a Cronbach's alpha greater than 0.7 was considered acceptable [28]. Additionally, the average inter-item correlation (AIC) and McDonald's omega (ω) were examined. The formula for calculating McDonald's omega is as follows: Omega = 1-[(N-SUM h2)÷*(N + 2r)] where N is the number of items, Sum h2* is the sum of item communalities, and r is the sum of the factor loadings of the study items. According to Glen [29], the desirable range for AIC is between 0.15 and 0.5, and McDonald's omega should be greater than 0.7 [30].

Test-retest method

In this study, the test-retest method was used to assess the stability of the scale. The scale was administered to 30 nurses in two stages, two weeks apart. Then, the Intraclass Correlation Coefficient (ICC) was calculated using a two-way mixed model with absolute agreement [31]. In this study, an ICC greater than 0.75 was considered acceptable. Moreover, to calculate the required sample size, a power analysis was conducted with alpha = 0.05 and power = 0.8 using the PASS software (version 11.0.7; PASS, NCSS, LLC) [32]. Based on these criteria, the required sample size was determined to be 30. The Intraclass Correlation Coefficient (ICC) indicates the relative reliability of the scale [21], and to assess absolute reliability, in addition to the ICC, the Standard Error of Measurement (SEM) was also calculated. The formula for calculating SEM is as follows:

$$SEM = SD_{Pooled} \times \sqrt{(1 - ICC_{agreement})}$$

In this formula, SD pooled is calculated using the equation $SD_{Pooled} = (SD1 + SD2)/2$.

Results

Item development and initial scale design

Based on an inductive-deductive approach, an initial pool of 73 items was developed. Subsequently, through a review session conducted by the research team, overlapping items were merged, and irrelevant items were eliminated. The final version of the scale comprised 37 items.

Scale validation

Content validity

Face validity

Quantitative assessment of face validity In this study, six items received a rating of less than 1.5 (on a range from 0 to 4.8) and were deemed inappropriate by respondents (Appendix 1).

Qualitative assessment of face validity Based on interviews with 10 nurses, the six items identified in the quantitative phase of face validity were revised (Appendix 2). The goal of this phase was to improve the clarity and comprehension of the items for the tool's users (nurses) by modifying the wording of ambiguous items.

Content validity

Qualitative assessment of content validity After interviewing 12 faculty members and considering their feedback, all items were revised (Appendix 3).

Quantitative assessment of content validity

Determining the content validity ratio (CVR) Following the qualitative content validity phase, the scale entered the CVR phase. In this phase, 21 items were removed due

to a score below 0.56, leaving the scale with 16 items for CVI analysis (Appendix 4).

Determining the content validity index (CVI) Based on the K* statistic, no items were removed at this stage, and the scale with 16 items proceeded to the initial reliability assessment (Appendix 5).

Initial reliability assessment

Following the establishment of content validity, the initial reliability of the scale was examined. In this stage, item analysis was conducted to obtain Cronbach's alpha and identify items that might affect the scale's reliability. To this end, 30 nurses who were directly involved in patient care at the hospital completed the scale. The initial results indicated a Cronbach's alpha of 0.594, which was not considered ideal. Subsequently, the Corrected Item Total Correlation was examined, and 4 items with a correlation of less than 0.1 were removed. After removing these items, the Cronbach's alpha of the scale increased to 0.883, which is considered desirable and acceptable. Ultimately, the scale with 12 items entered the construct validity phase.

Construct validity

To determine the construct validity, exploratory factor analysis (EFA) was employed. Initially, 225 nurses participated in the study. However, after excluding indifferent participants and univariate and multivariate outliers, the sample size was reduced to 222.

The mean and standard deviation of the participants' age were 31.60 ± 7.93 years. Out of the 222 participants, 72 (32.4%) were male and 150 (67.6%) were female. Additionally, 99 (44.6%) of the participants were married and 123 (55.4%) were single. The participants had an average work experience of 7.55 ± 7.41 years. Regarding their education, 169 (76.1%) participants held a bachelor's degree, and 53 (23.9%) held a master's degree. The participants were from various departments: 64 (28.8%) from the internal medicine department, 26 (11.7%) from the surgery department, 42 (18.9%) from the emergency department, 19 (8.6%) from the ophthalmology department, 10 (4.5%) from the ENT department, and 61 (27.5%) from the special care unit. Regarding their work shifts, 25 (11.3%) worked day shifts, 5 (2.3%) worked night shifts, and 192 (86.5%) worked rotating shifts.

To assess the adequacy of the sample size, the Kaiser-Meyer-Olkin (KMO) measure was estimated to be 0.834, which is considered excellent. Moreover, Bartlett's test of sphericity was significant ($P < 0.001$). In determining the number of factors, both the scree plot and parallel analysis supported the extraction of two factors. The scree plot indicated two potential factors, as only the first two eigenvalues exceeded the 95th percentile of simulated data (Fig. 1). Similarly, parallel analysis, presented in Table 4, demonstrated that the observed eigenvalues for the first two factors were greater than the corresponding random eigenvalues, confirming their significance and justifying the retention of two factors. Also, based on the 95% confidence interval in the Table 4, it can be confirmed with 95% confidence that the eigenvalue for both

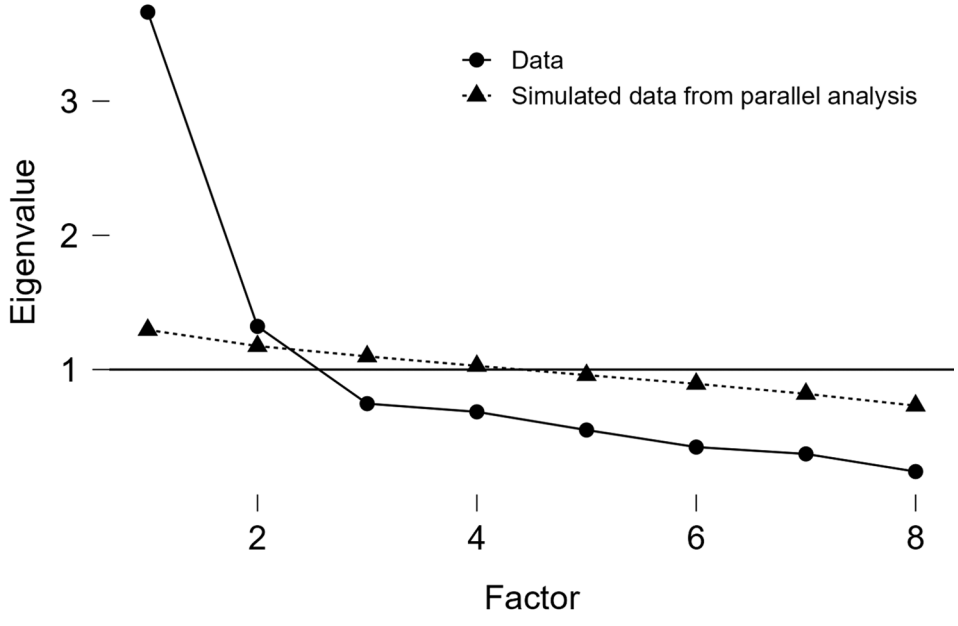

Fig. 1 Scree plot used in exploratory factor analysis to visually determine the number of underlying factors of the honesty scale for nurses

Table 4 Parallel analysis results based on observed and random eigenvalues

Factor	Observed eigenvalue (95% confidence interval)	Random eigenvalue
1	3.663 (3.315–4.011)	1.314
2	1.322 (1.196–1.447)	1.205
3	**0.746 (0.691–0.836)**	**1.126**
4	0.685 (0.619–0.751)	1.057
5	0.550 (0.497–0.602)	0.925
6	0.423 (0.382–0.463)	0.862

extracted factors is equal to or greater than one. After determining the three extracted factors and performing the final analysis, the number of items in the scale was reduced to 8. Based on Table 5, the final scale explained 51.034% of the total variance.

The minimum acceptable factor loading for conducting exploratory factor analysis was calculated to be approximately 0.35 using the formula. Therefore, the factor loadings extracted after varimax rotation were reported. Accordingly, based on the varimax rotation, the highest extracted factor loading was 0.877 and the lowest was 0.578. It is worth noting that the factor loading of all items was above 0.5 (Table 6).

Factor naming

After extracting the factors, each factor was named based on its constituent items, as detailed in Table 7.

Reliability assessment

Internal consistency

In this study, Cronbach's alpha coefficient was calculated to determine the internal consistency of the scale. Additionally, considering that Cronbach's alpha increases artificially with increasing sample size and number of items [21], McDonald's omega (ω) was also reported. Accordingly, Cronbach's alpha for the entire scale was reported as 0.823, and McDonald's omega for the entire scale was 0.898. Also, the AIC in this study was 0.347.

Test-retest reliability

In this study, the test-retest method was used to determine the stability of the scale. To this end, the scale was completed by 30 nurses and then, after two weeks, the same participants completed the scale again. The intraclass correlation coefficient (ICC) for the entire scale between the two measurements was reported as 0.918, with a 95% confidence interval of 0.827 to 0.961; also, the standard error of measurement for the entire scale was calculated as 1.246.

The final version of the scale is in Appendix 6.

Table 5 Eigenvalues and explained variance for the extracted factors

Factor	Initial Eigenvalues			The sum of the squares of the extracted factor loadings			The sum of the squares of the extracted factor loadings		
	Total	Variance	The cumulative percentage	Total	Variance	The cumulative percentage	Total	Variance	The cumulative percentage
1	3.663	45.784	45.784	3.037	37.961	37.961	2.084	26.053	26.053
2	1.322	16.524	62.308	1.046	13.072	51.034	1.998	24.981	51.034
3	0.746	9.328	71.636						
4	0.685	8.564	80.200						
5	0.550	6.870	87.071						
6	0.423	5.282	92.352						

Table 6 Extracted factor loadings after varimax rotation

Item number	Items	Factor loading	
		Factor 1	Factor 2
1	I am an honest person in my job and always tell the truth to my patients.		0.669
2	I answer patients' questions honestly.		0.877
3	If an action for the patient is not performed during my shift, I honestly report it to the next shift staff.	0.662	
4	I take responsibility for my individual errors and promptly take actions to correct and compensate for them.	0.620	
5	To improve communication, I avoid lying and giving incorrect information to my patient.		0.768
6	I document the complete details of the care and treatment actions performed for my patient accurately and honestly.	0.599	
7	In case of an allergic reaction in the patient due to my individual error, I promptly inform the doctor and other colleagues and immediately start treatment actions for the patient.	0.578	
8	I perform patient care and treatment actions with precision, patience, and accuracy.	0.628	

Table 7 Factor labeling and description

Factor's name	Factor description
The first factor: **Behavioral honesty**	Behavioral honesty refers to an individual's adherence to ethical values and commitment to acting in accordance with them. It encompasses aspects such as honest reporting, accountability, accurate documentation, prompt response to emergencies, and diligent task completion
The second factor: **Verbal honesty**	Verbal honesty denotes truthfulness and transparency in communication. This includes elements such as honesty in speech, providing truthful answers to questions, and refraining from deceiving others through language.

Discussion

The developed questionnaire, as an innovative tool, offers an integrated approach to assessing honesty in nursing by combining dimensions of behavioral and verbal honesty. Unlike existing instruments [14–16], which often focus on ethical sensitivities or related concepts, this questionnaire comprehensively captures honesty as a distinct and measurable construct. By incorporating detailed psychometric evaluations, the tool addresses nuances such as context-specific honesty behaviors and truthful communication, both critical to nursing practice. Its comprehensive design ensures it can serve as a benchmark for evaluating and promoting ethical standards across various healthcare settings. This highlights its potential to fill the gaps identified in previous studies, reinforcing the need for such a dedicated instrument in nursing.

The first dimension of the questionnaire is behavioral honesty. This dimension accounts for 26.05% of the total variance in the study. In this questionnaire, behavioral honesty is a conceptual explanatory variable where an individual adheres to ethical values and is committed to practicing them. This factor includes aspects such as honest reporting, acceptance of responsibility, accurate documentation, quick response in emergencies, and accurate performance of duties. In this scale, after conducting psychometric procedures on the items developed in the item pool stage, the items measuring the level of nurses' behavioral honesty were assessed in the form of five items with the general concept of nurses' responsibility and accuracy in patient care in various conditions.

In the scale designed by Muramatsu et al. items with themes such as providing necessary care despite patient resistance due to medical necessity (e.g., changing a patient's position post-surgery) and providing long-term care with patience (e.g., assisting a dysphagic patient with feeding) were consistent with the concept of providing correct and patient care in the present scale. Additionally, items related to accurately reporting the patient's condition to the nurse in charge, emphasizing honesty in reporting, were aligned with other items in the scale in this area. However, the present scale covers various dimensions of honest reporting and accurate task performance more comprehensively than the questionnaire by Muramatsu et al. [14].

Furthermore, the statement "I think of every action so that I can give an account", which was included in the responsibility dimension of the scale designed by Suryadi et al., can be aligned with the acceptance of responsibility in the current scale [33]. Additionally, the component of accurate documentation in this dimension is reflected in a statement such as "Writing up your written assignment when you have done the work" in the scale by Henning et al. which expresses honesty in documentation [34]. However, none of the scales used in the field of behavioral honesty have addressed the issue of rapid response in emergencies; moreover, most of the existing scales have not covered aspects of bedside care in their items.

The second dimension of the current scale is verbal honesty. This dimension accounts for 24.981% of the total variance of the scale. Verbal honesty in this questionnaire explains the concept of truthfulness and transparency in communication. This factor includes aspects such as truthfulness in speech, honest answers to questions, and refraining from deceiving others through speech. The items present in the verbal honesty dimension, similar to the behavioral honesty dimension, were reduced to three items after completing the psychometric stages and modifying and removing inconsistent items. The items present in this dimension include "I am an honest person in my job and always tell the truth to my patients," "I answer patients' questions honestly," and "To improve communication, I avoid lying and giving incorrect information to my patient".

The scales by Muramatsu et al. [14] and Shabir et al. [35] include concepts similar to those in the verbal honesty dimension. However, these scales differ in their target population (nursing students in Muramatsu et al.; leadership integrity in Shabir et al.). Additionally, they do not comprehensively cover all aspects of verbal honesty in nurses, such as providing honest answers to patients' questions [14, 35].

To design and psychometrically assess the nursing honesty questionnaire, after forming the item pool, the following steps were taken in order: face validity, content validity, initial reliability, construct validity, and final reliability.

In determining the face validity of the current questionnaire, the opinions of nurses, as the future users of the scale, were used. In reviewing scales with items similar to those designed in the present study, it was found that face validity was only reported in the scale designed by Muramatsu et al., and no evidence was found in the

review of other studies regarding the determination of face validity [14]. This is while determining face validity and whether the items of each dimension are logical, appropriate, and relevant for those who routinely use this tool as a criterion is of great importance [36].

Furthermore, face validity is a reflection of the clarity or relevance of an scale for both respondents and examiners, and it can affect the quality of responses to questions [37, 38]. In the present study, the written structure of the items was made clear and understandable for the target audience (nurses) using a qualitative method, and ambiguities in the writing style were eliminated as much as possible. Also, after determining the face validity using qualitative methods, the quantitative method showed that all items with a coefficient above 1.5 were considered appropriate by the participants. Thus, according to Ronald Holden, the modification and revision of items in face validity reduced the gap between the scale and the respondent as much as possible [39].

In the present study, content validity was assessed using both qualitative and quantitative methods, and expert opinions were obtained. Conducting content validity in scale development studies is directly related to gaining confidence in the instrument's ability to measure the concept under study (honesty in nursing) and can corroborate researchers' claims about the current instrument's measurement capabilities. Among the studies conducted on scales measuring the concept of honesty, the "Personal Integrity Scale" designed by Suryadi et al., which considered integrity as equivalent to honesty, did not assess content validity [33]. This is while the absence of content validity can be considered one of the major shortcomings of instrument development studies. Content validity plays a significant role in scale development, and its calculation and indices are crucial, especially when the scale is used to measure health outcomes or guide clinical decision-making [40]. On the other hand, the lack of reporting content validity also questions the desirability of the reliability of that scale [41].

In this study, construct validity was assessed using exploratory factor analysis. Consequently, the sample size was evaluated at a high level. According to Horn, parallel analysis is one of the best indicators for determining the number of factors needed for extraction [42]. Therefore, in addition to classical criteria, novel indices such as parallel analysis were used. Interpretation and analysis of factor analysis indicated that the final scale in the present study, with 8 items and in two dimensions, explained 51.03% of the total variance, which is considered acceptable considering the 50% criterion. In reviewing the methodology of existing scales related to the concept of honesty, despite the fact that factor analysis is considered a powerful technique in determining construct validity and discovering the underlying variables of a phenomenon, no clear and detailed report of factor analysis was found in the "Personal Integrity Scale" scale [33].

In the present study, internal consistency was reported as 0.823 for the entire scale using Cronbach's alpha, which was evaluated at a satisfactory level. In addition, McDonald's omega coefficient was used in this study to estimate the reliability of the scale more accurately. The numerical value of this coefficient for the entire scale was 0.898. Also, to assess the stability of the scale, the intraclass correlation coefficient (ICC) for the entire scale was 0.918, with a 95% confidence interval of 0.827 to 0.961, indicating a satisfactory level of reliability over time for this study. Consistent with this finding, the "Ethical Sensitivity Questionnaire for Nursing Students" also examined the stability of the scale through test-retest with an eight-week interval and reported a Pearson correlation coefficient of 0.42, which, considering that the desirable level is reported to be 0.71 and above, does not report very ideal conditions for the "Ethical Sensitivity Questionnaire for Nursing Students" and is in the moderate range. On the other hand, in this scale, no report was found regarding the ICC [14]. In the present study, following the recommendation of Dutil et al., a two-week interval between the two tests was considered appropriate [43]. In other studied scales, including the "Personal Integrity Scale" scale, test-retest and ICC reporting were not mentioned [33]. Another scale, the "Perceived Leaders' Integrity Assessment Scale", although the stability of the scale was not measured using the test-retest method, reported an alpha reliability coefficient of 0.93, indicating high reliability and internal consistency [35].

Another important point is that none of the scales used to measure the concept of honesty have made any attempt to determine absolute reliability. In other words, indices such as SEM, along with ICC, would provide absolute reliability and indicate whether the difference in measurement between two tests is real or due to measurement error, were not reported. In the present study, SEM was estimated to be 1.246 for the entire scale. The low level of SEM and the high level of ICC in this scale indicate its high accuracy and the strength of its absolute reliability.

Overall, a review of scales used to measure the concept of honesty revealed that four main scales were used, including the "Ethical Sensitivity Questionnaire for Nursing Students", "Personal Integrity Scale", "Perceived Leaders' Integrity Assessment Scale", and "Cross-Cultural Academic Integrity Questionnaire". To critique each of these scales, as mentioned, methodological flaws were identified in the psychometric sections and stages of these studies, all of which reduced the efficiency and accuracy of measuring the concept of honesty. Moreover, the aforementioned scales were used in different target populations, which, for all the reasons mentioned, this

study has been able to develop a high-quality scale, in accordance with scale development standards, that has to a large extent addressed the shortcomings of previous studies and has been able to provide a suitable scale for conducting necessary studies by other researchers.

This newly developed nursing honesty questionnaire, with its focus on behavioral and verbal honesty, offers significant practical implications for nursing practice. By providing a reliable and valid tool to assess these crucial aspects of professional conduct, healthcare institutions can utilize the questionnaire for various purposes. For example, it can be employed in pre-employment screening to identify candidates with strong ethical foundations. Furthermore, it can serve as a valuable instrument for ongoing professional development, allowing nurses to self-assess and identify areas for improvement in their honesty and integrity. The questionnaire can also be integrated into performance evaluations, offering a structured approach to discussing and reinforcing ethical behavior. Ultimately, by fostering a culture of honesty and transparency, this tool can contribute to enhanced patient safety, improved interprofessional communication, and increased public trust in the nursing profession.

One limitation of the present study is the absence of confirmatory factor analysis (CFA). While exploratory factor analysis (EFA) was utilized to identify the underlying factors of the scale, CFA could have provided further validation of the factor structure and its fit to the data. Therefore, it is recommended that future studies employ CFA to confirm the factor structure derived from the EFA and to further establish the construct validity of the honesty scale among nurses.

Conclusion

This study resulted in the development of an 8-item scale with two dimensions: behavioral and verbal honesty. The developed scale, in accordance with the standards of scale development studies, has the necessary validity to measure the level of honesty in nursing. With acceptable values of indices such as validity and reliability, it has been able to significantly reduce the existing research gap in this area. Thus, a scale was developed that enables the measurement of the level of honesty in nurses, where higher scores indicate higher levels of honesty and lower scores indicate lower levels of honesty among nurses. Future research should focus on cross-cultural validation of this tool across diverse healthcare settings and countries to establish its universal applicability. In clinical practice, we recommend that nursing managers incorporate this tool in performance evaluations and professional development programs, creating targeted interventions for departments with lower honesty scores. Educational institutions could utilize this scale to assess and reinforce ethical conduct among nursing students, potentially integrating honesty measurement into curriculum development. Healthcare organizations might consider implementing regular honesty assessments as part of quality improvement initiatives, establishing benchmarks and recognizing units demonstrating high ethical standards, which could ultimately contribute to improved patient care outcomes and organizational culture1. Additionally, the scale could be adapted for specialized nursing contexts such as intensive care, emergency departments, or community settings where unique ethical challenges may arise.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12912-025-03163-0.

Appendix 1. The item impact score in the quantitative face validity.

Appendix 2. Revised items in the qualitative Face validity phase.

Appendix 3: Revised items in the qualitative content validity stage.

Appendix 4. CVR results of items in the quantitative content validity.

Appendix 5: I-CVI results based on K*.

Appendix 6: The final version of the scale.

Acknowledgements

We would like to express our sincere gratitude to the Student Research Committee at Kermanshah University of Medical Sciences, Kermanshah, Iran, for their financial support (no 4020656).

Author contributions

MMM, AN, HM, and MM conceptualized, analyzed and interpreted the data, and drafted the manuscript. MMM and MM designed the study and participated in the analysis and interpretation of data. All authors coordinated the study, revised the manuscript, edited and approved the final version to be submitted for publication, and helped in the analysis and interpretation of data.

Funding

The present study was financially supported by Kermanshah University of Medical Sciences, Kermanshah, Iran (no. 4020656). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Data availability

The data analyzed and materials used in this study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

In accordance with the Declaration of Helsinki and principles of ethical research, this study was approved by the Ethics Committee of Kermanshah University of Medical Sciences (Ethics code: IR.KUMS.REC.1402.350). Participants were assured of the confidentiality of their data. They were also informed of their right to withdraw from the study at any time. Informed consent was obtained from all individual participants included in the study.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 24 December 2024 / Accepted: 5 May 2025
Published online: 14 May 2025

References

1. Park N, Peterson C, Seligman ME. Character strengths in Fifty-four nations and the Fifty US States. J Posit Psychol. 2006;1(3):118–29.
2. Błachnio A, Be Happy B, Honest. The role of Self-Control, Self-Beliefs, and satisfaction with life in honest behavior. J Relig Health. 2021;60(2):1015–28.
3. Lenton A, Bruder M, Slabu L, Sedikides C. How does being real feel? The experience of state authenticity. J Pers. 2013;81:276–89.
4. Eijkholt M, Sparling A. Health, honesty and happiness: authenticity and anonymity in social media participation of individuals with multiple sclerosis. Multiple Scler Relat Disorders. 2019;27:121–6.
5. Vivekananda N, Meenakshi DR. Honesty and its role in maintaining social cohesion and trust. Educational Adm Theory Practices. 2024;30.
6. Motivation Hub IBEKWEO, Musings M. Thinking for A Living: Osita IBEKWE 2024. Available from: https://ositaibekwe.com/reasons-why-honesty-is-not-the-best-policy/
7. Shalvi S, Levine E, Thielmann I, Jayawickreme E, van Rooij B, Teodorescu K et al. The Science of Honesty: A Review and Research Agenda. 2024.
8. Ichsan RN, Nasution L, Setiadi D. The influence of work ethics and work professionalism on performance at Pt. Bri branch Singamangaraja Medan. Jurnal Darma Agung. 2022;30(1):118–25.
9. Panigrahi S, Al-Nashash HM. Quality work ethics and job satisfaction: an empirical analysis. Qual Access Success. 2019;20.
10. Fitriah M, Madjid A, Honesty. A multidimensional study as motivation for National character Building. Hayula: Indonesian J Multidisciplinary Islamic Stud. 2020;4(1):99–116.
11. Sastrawan S, Newton JM, Malik G. Nurses' integrity and coping strategies: an integrative review. J Clin Nurs. 2019;28(5–6):733–44.
12. Markey K, Moloney M, Doody O, Robinson S. Time to re-envisage integrity among nurse leaders. J Nurs Adm Manag. 2022;30(7):2236–40.
13. Devine CA, Chin ED. Integrity in nursing students: A concept analysis. Nurse Educ Today. 2018;60:133–8.
14. Muramatsu T, Nakamura M, Okada E, Katayama H, Ojima T. The development and validation of the ethical sensitivity questionnaire for nursing students. BMC Med Educ. 2019;19(1):1–8.
15. Wote AYV, Sero WP. The contribution of principals' leadership role and teachers' integrity to elementary school teachers' performance. Int J Elementary Educ. 2021;5(2):330–6.
16. Henning MA, Nejadghanbar H, Abaraogu U. Developing a revised cross-cultural academic integrity questionnaire (CCAIQ-2). J Acad Ethics. 2018;16:241–55.
17. Boateng GO, Neilands TB, Frongillo EA, Melgar-Quinonez HR, Young SL. Best practices for developing and validating scales for health, social, and behavioral research: A primer. Front Public Health. 2018;6:149.
18. Avila ML, Stinson J, Kiss A, Brandao LR, Uleryk E, Feldman BM. A critical review of scoring options for clinical measurement tools. BMC Res Notes. 2015;8:612.
19. Colton D, Covert RW. Designing and constructing instruments for social research and evaluation. Wiley; 2007.
20. Polit DF, Beck CT. Nursing research: generating and assessing evidence for nursing practice. Lippincott Williams & Wilkins; 2017.
21. Polit D, Yang F. Measurement and the measurement of change. Wolters Kluwer; 2016.
22. Polit DF, Beck CT, Owen SV. Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. Res Nurs Health. 2007;30(4):459–67.
23. Munro BH. Statistical methods for health care research. lippincott williams & wilkins; 2012.
24. Guadagnoli E, Velicer WF. Relation of sample size to the stability of component patterns. Psychol Bull. 1988;103(2):265.
25. Larsen R, Warne RT. Estimating confidence intervals for eigenvalues in exploratory factor analysis. Behav Res Methods. 2010;42(3):871–6.
26. Finch H, French BF, Immekus JC. Applied psychometrics using SPSS and AMOS: Information Age Publishing 2016.
27. Norman GR, Streiner DL. Biostatistics: the bare essentials. 4th Edition ed: PMPH USA; 2014.
28. Taber KS. The use of Cronbach's alpha when developing and reporting research instruments in science education. Res Sci Educ. 2018;48(6):1273–96.
29. Glen S. Average Inter-Item Correlation: Definition, Example From StatisticsHowTo.com: Elementary Statistics for the rest of us! 2021. Available from: https://www.statisticshowto.com/average-inter-item-correlation/
30. Stone LL, Janssens JM, Vermulst AA, Van Der Maten M, Engels RC, Otten R. The strengths and difficulties questionnaire: psychometric properties of the parent and teacher version in children aged 4–7. BMC Psychol. 2015;3(1):4.
31. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med. 2016;15(2):155–63.
32. Bujang M, Baharum N, editors. A simplified guide to determination of sample size requirements for estimating the value of intraclass correlation coefficient: a review2017.
33. Suryadi B, Nisa Y, Sumiati N. Development of personal integrity scale: Construct Validity2018.
34. Henning M, Alyami M, Melyani Z, Alyami H, Al-Mansour A. Development of the Cross-Cultural academic integrity Questionnaire - Version 3 (CCAIQ-3). J Acad Ethics. 2020;18:35–53.
35. Muazzam PDN, Koppel A, Shabir S. S. Development and Validation of the Perceived Leaders' Integrity Assessment Scale. Pakistan J Humanit Social Sci. 2024;12.
36. Connell J, Carlton J, Grundy A, Taylor Buck E, Keetharuth AD, Ricketts T, et al. The importance of content and face validity in instrument development: lessons learnt from service users when developing the recovering quality of life measure (ReQoL). Qual Life Research: Int J Qual Life Aspects Treat Care Rehabilitation. 2018;27(7):1893–902.
37. Ebadi A, Zarshenas L, Rakhshan M, Zareiyan A, Sharif Nia H, Mojahedi M. Principles of Scale Development in Health Science2017.
38. Paulisso DC, Cruz DM, Allegretti ALC, Schein RM, Costa JD, Campos LC, et al. Cross-Cultural adaptation and face validity of the functional mobility assessment into Brazilian Portuguese. Occup Therapy Int. 2020;2020(1):8150718.
39. Holden RB. Face Validity. The Corsini Encyclopedia of Psychology. 2010:637-8.
40. Almanasreh E, Moles R, Chen TF. Evaluation of methods used for estimating content validity. Res Social Administrative Pharm. 2019;15(2):214–21.
41. Rubio DM, Berg-Weger M, Tebb SS, Lee ES, Rauch S. Objectifying content validity: conducting a content validity study in social work research. Social Work Res. 2003;27(2):94–104.
42. Horn JL. A rationale and test for the number of factors in factor analysis. Psychometrika. 1965;30:179–85.
43. Dutil E, Bottari C, Auger C. Test-Retest reliability of a measure of independence in everyday activities: the ADL profile. Occup Ther Int. 2017;2017:3014579.

Publisher's note