Journal of Transcultural Nursing
2025, Vol. 36(5) 480–491
© The Author(s) 2025
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/10436596251330295
journals.sagepub.com/home/tcn
Sage

Development and Psychometric Testing of the Cultural Bias Scale for Health Professionals: A Methodological Study

Esin Kavuran, PhD[1], Afife Yurttaş, PhD[1], Tuğçe Kabak Solak, MS[1], and Üzeyir Solak, MS[1]

Abstract

Introduction: Cultural bias can negatively affect processes such as making the correct diagnosis, creating appropriate treatment plans, and ensuring patient satisfaction when providing health care services to individuals. This study aims to develop and conduct the psychometric evaluation of a reliable tool to measure the bias levels of health care professionals. **Method:** The scale was developed with a pilot study conducted on 50 nurses in Erzurum, Turkey. The scale initially consisted of 53 items, and data collection was performed on 720 participants. **Results:** The scale Kaiser–Meyer–Olkin (KMO) value was found to be 0.861. The anti-image value for items was ≥0.5. Bartlett's sphericity tests indicated suitability for factor analysis ($\chi^2 = 2{,}850.732$, $df = 210$, $p < .01$). The variance explained by each subscale of the scale was 18.582%, 16.342%, and 14.507%, respectively. **Discussion:** Because health care professionals should take a holistic approach to patient care, the scale might be a useful a important instrument for evaluating the cultural biases of health care professionals.

Keywords

cultural bias, ethnonursing, health care professionals, scale development, culturally competent

Introduction

With globalization, many people are compelled to migrate, either willingly or unwillingly, due to factors such as war, regional conflicts, and poverty. Migration refers to the movement of individuals or groups driven by economic, cultural, or social reasons (De Haas et al., 2019). This situation has led to individuals with diverse languages, religions, cultures, health beliefs, and life experiences receiving care from different health care providers (Kirmayer & Jarvis, 2019). Cultural differences play a significant role in nursing, as health care professionals cater to individuals from various ethnic and cultural backgrounds. Therefore, it is critical for health care professionals to understand and acknowledge these cultural nuances to provide effective and inclusive health care services (Baumgarten, 2018). Ensuring that health care services are accessible, acceptable, and effective for everyone requires nurses to possess cultural competence (Almutairi et al., 2017).

Various scales have been developed in the literature to assess cultural competence and related constructs in health care professionals. For example, the "Cultural Competence Assessment Instrument" (Campinha-Bacote, 2002) evaluates health care workers' cultural competence skills, while the "Cultural Awareness Scale" (Rew et al., 2003) focuses on measuring nurses' levels of cultural awareness (Campinha-Bacote, 2002; Rew et al., 2003). However, these tools fail to fully capture the nuances of cultural bias within health care settings, especially in multicultural societies such as Turkey. While these tools contribute to understanding broader concepts like cultural competence, they do not specifically address the biases and discrimination that may occur within health care environments.

Nursing is a profession aimed at protecting and promoting the health of individuals, families, and communities and at restoring and maintaining their physical, mental, and social well-being when their integrity is compromised (Sastrawan et al., 2019). Nurses are among the professions that spend the most time and the longest duration with individuals and families of diverse cultures, beliefs, and attitudes during the provision of care (Duffy, 2022). The International Council of Nurses (ICN) Code of Ethics (2019) states that "nursing is a universal need, and at its core,

[1]Ataturk University, Erzurum, Turkey

Corresponding Author:
Esin Kavuran, Assistant Professor, Department of Nursing Fundamentals, Faculty of Nursing, Ataturk University, 25240 Erzurum, Turkey.
Email: esin.kavuran@atauni.edu.tr

nursing values and respects human life and human rights without discrimination based on nationality, language, religion, gender, belief, age, political opinions, or social status" (Tisdale & Symenuk, 2020). To provide comprehensive care, nurses must deeply understand an individual's cultural heritage, values, and potential biases. Nurses must also be aware of their own backgrounds, experiences, values, and biases, and how these factors might influence their interactions with patients (Kaya & Boz, 2019).

Culturally competent nursing not only promotes positive patient outcomes but also strengthens trust, communication, and mutual understanding between health care providers and patients (Tang et al., 2019). Due to its geographical location, Turkey is defined as a "multidimensional migration country." It is crucial for health care professionals in Turkey to be aware of the various cultural landscapes that arise from the country's geopolitical position (Karatas, 2023). In the presence of multiple cultures within a societal structure, it is essential for health care providers to have a high level of cultural awareness (Baumgarten, 2018). Assessing nurses' perceptions of bias and discrimination within the health care system is vital to identifying key factors contributing to health outcome inequalities (Durand et al., 2022). In this context, Kressin et al. (2008) suggest the development of assessment tools that measure perceptions of discrimination toward different racial and ethnic groups within health care environments. Such tools will enable health care professionals to better understand and address potential biases, ensuring fair and inclusive care for all individuals (Kressin et al., 2008).

It is necessary to develop a tool grounded in a theoretical framework to address the effects of cultural bias in the provision of health care services. In this regard, Leininger's Transcultural Nursing Theory has been adopted as a guiding framework. Leininger argues that cultural sensitivity and competence play a crucial role in maintaining the health of individuals and communities. This theory requires awareness of cultural beliefs and values, which influence how individuals perceive health, illness, and care. Nurses must be sensitive to the cultural backgrounds of the individuals they serve and avoid bias. According to the theory, cultural bias may arise when a health care worker approaches a patient from the perspective of their own culture's superiority (Leyva-Moral & Bernabeu-Tamayo, 2021). Therefore, health care professionals must be aware of their biases and control them. Leininger's theory promotes the development of cultural awareness and empathy, enabling health care professionals to communicate better with individuals from diverse cultural backgrounds (Lima et al., 2023). Despite the existence of numerous scales measuring general cultural competence and awareness, there is a significant gap in tools specifically designed to assess cultural bias within health care settings. This study aims to fill this gap by developing a new measurement tool focused on assessing cultural bias among health care providers. The objective of this study is to develop a tool that can effectively assess health care providers' levels of cultural bias in health care environments. The absence of an existing tool in the literature for measuring cultural bias highlights the importance and urgency of this research. The aim of this study is to develop a valid and reliable scale to measure cultural bias in clinical nurses and to assess its psychometric properties. By doing so, we aim to enhance the efficiency of health care services by minimizing unnecessary and inappropriate interventions.

Method

The study comprised two phases: (a) the creation of the item pool and (b) the psychometric evaluation of the scale, following the guidelines in the literature (Boateng et al., 2018; Kyriazos & Stalikas, 2018). The flowchart is displayed in Figure 1 and the reporting process adhered to the STROBE criteria.

Creating the Item Pool

The researchers clearly defined the measurement objective and created an item pool (Kyriazos & Stalikas, 2018). Between 2000 and 2022, a comprehensive literature search was conducted using databases such as PubMed, Web of Science, CINAHL, Science Direct, Scopus, and the National Index. Search terms included combinations of (nurse OR nursing) AND (cultural competence OR cultural awareness), (cultural diversity OR cultural sensitivity) AND (healthcare provider OR nursing), (cultural bias OR implicit bias) AND (healthcare OR nursing practice), and (transcultural nursing OR intercultural competence) AND (patient care OR clinical outcomes). As a result of the literature search, 245 articles were identified, of which 45 were excluded due to lack of full-text access, 30 were duplicates, and 25 were unrelated to the topic. Therefore, a total of 145 articles were included for further analysis. Inclusion criteria were as follows: (a) studies focused on cultural competence and bias in nursing and health care, (b) full-text availability, (c) published between 2000 and 2022, and (d) published in peer-reviewed journals. Articles that were available only as abstracts or were unrelated to the topic were excluded. Data were extracted from the included articles, focusing on the study's objectives, methodology, findings, and how items related to cultural bias were developed. In particular, the item development processes used in the articles and how the content of these items was assessed were analyzed. The quality of the articles was evaluated based on criteria such as the robustness of the research methodology, adequacy of sample size, appropriateness of data analysis techniques, and consistency of the findings. Specific criteria included the validity of the methodological approach, reliability of the findings, and the overall suitability of the research design (Boateng et al., 2018) . During the literature review, existing studies on cultural bias, cultural competence, and

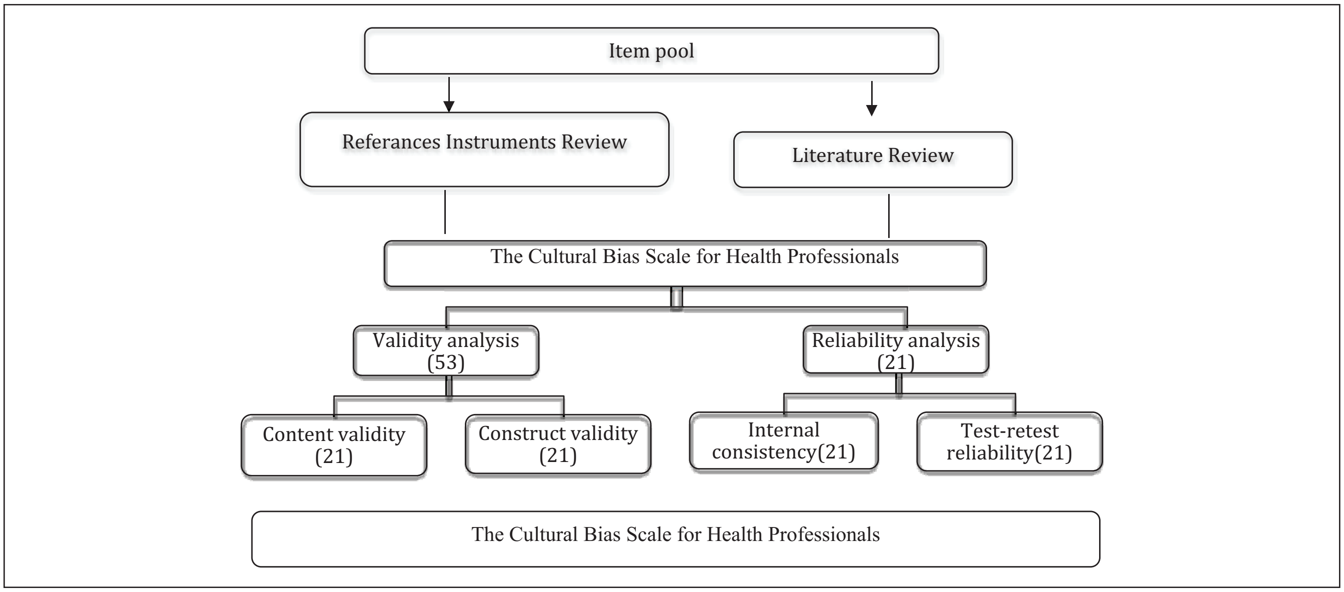

Figure 1. STROBE Guidelines for the Cultural Bias Scale for Health Professionals.

awareness were carefully analyzed to extract key findings. These findings included issues such as the biases faced by health care professionals when working with patients from different ethnic and socioeconomic backgrounds, communication difficulties, and their impact on the quality of care. The themes that emerged from this analysis were used as the foundation for developing the subscales of the instrument. The Care Process Factors subscale was developed based on studies that addressed the effects of cultural differences on patient care as described in the literature. These items were formulated to assess the cultural barriers health care professionals face during care and how these barriers reflect on care processes. The Individual Opinion Factors of the Health Professional subscale was shaped by reviewing findings in the literature that focused on health care professionals' personal biases and levels of cultural sensitivity. The items in this subscale were designed to evaluate the impact of professionals' personal biases on patient care. The Socioeconomic Factors in Patient Care subscale was developed based on findings from the literature addressing the effects of patients' socioeconomic status on access to health care and care processes. These items were created to measure how health care professionals perceive patients' economic and social conditions and how these perceptions influence the quality of care. Key concepts highlighted in the literature (e.g., cultural sensitivity, patient safety, quality of care) played a critical role in the development of the scale items. Each item was formulated as a concrete reflection of these themes, and content validity was ensured through expert feedback.

Design, Sample, and Setting. This study employed a methodological approach within a cross-sectional study design to develop a valid and reliable measurement tool to assess cultural bias in clinical nurses. The methodological approach adhered to established guidelines for scale development and validation and involved specific steps, including item generation through literature review, expert validation for content relevance, pilot testing to evaluate clarity and feasibility, and psychometric evaluation to establish reliability and validity. First, a literature review was conducted to determine the scope of the instrument, and then expert opinions were obtained to ensure content validity. Pilot studies were conducted to test the clarity and validity of the scale items, and in the final stage, the construct validity of the scale was evaluated using statistical methods such as factor analysis. This process was carefully implemented to ensure that the measurement tool is both valid and reliable. Data collection was carried out in two hospitals in Erzurum, Turkey, between February and June 2023. The data collection process was carried out between February and June 2023 in two hospitals located in Erzurum that is situated in the Eastern Anatolia Region of Turkey and has a unique socio-cultural structure that may reflect specific patterns of cultural bias observable across Turkey. In this context, the two hospitals selected for the study vary in terms of size, patient profiles, and types of services provided, creating a diverse sample that can reflect different levels of professional experience and biases among health care professionals. Sample sizes were calculated using a general rule of factor analysis that requires at least 10 participants for each item, with a 20% attrition rate (Sapnas & Zeller, 2002). The 53 items on the original scale required at least 636 participants. Data collection forms were therefore administered to 720 participants. In scale development, using separate data sets for exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) is crucial to validate the identified structure independently (Schumacker & Lomax, 2010). This can be achieved by randomly dividing a single data set into two groups. In this study, participants were randomly split: the odd group underwent EFA (n = 321) and the even group underwent CFA (n = 321). Incomplete surveys (n = 28) were excluded (Carpenter, 2018). In addition, on the recommendation of Bolarinwa (2015), a pilot group of 50 participants was not included in the main sample (Bolarinwa, 2015).

Psychometric Evaluation of the Scale

Content Validity. The content validity of the 53-item Cultural Bias Scale was assessed using the Davis technique (Davis, 1992). Eighteen doctoral-level nurses specializing in transcultural nursing were invited via email, and 14 (78%) participated. Experts were selected based on their knowledge of nursing, cultural competence, bias, and discrimination, ensuring diverse ethnic and cultural representation. In the Davis technique, experts evaluate each item on a 4-point Likert-type scale (1 = *not appropriate*, 4 = *completely appropriate*). The item-content validity index (I-CVI) was calculated by dividing the number of experts who rated an item as three or four by the total number of experts. The Scale-Content Validity Index (S-CVI) was computed as the mean of all I-CVI scores, with a minimum S-CVI of 0.90 and I-CVI of 0.78 required for inclusion (Almanasreh et al., 2019; Polit-O'Hara & Yang, 2016). Items with a corrected item-total correlation (CITC) below 0.20 were removed to improve the scale's reliability, and the study was rerun to enhance the Cronbach's alpha coefficient (Streiner et al., 2024).

Pilot Study. A sample group of at least 50 people with comparable features should be utilized to examine the content validity of scales (Yusoff et al., 2021). To improve the validity of the Cultural Bias in Health Professions scale and evaluate the items' appropriateness and comprehensibility, a pilot study with 50 nurses was carried out. The inclusion criteria for the pilot study required participants to have at least 2 years of clinical nursing experience and a minimum of a bachelor's degree in nursing, ensuring both practical experience and relevant knowledge. Nurses from various departments, such as internal medicine, surgery, intensive care, and emergency services, were included to provide a diverse perspective on how different work environments may influence cultural biases. Participants found the draft scale to be understandable and easy to read, but they suggested minor changes, such as making some expressions clearer and replacing certain terms with more commonly used words. No items were deleted, but participant feedback was incorporated. After receiving additional feedback from four care experts and a linguist, the final scale was prepared for the main sample.

Construct Validity. During the EFA process, several critical steps were undertaken to refine the scale's dimensions. First, the data were assessed for suitability using the Kaiser–Meyer–Olkin (KMO) measure and Bartlett's test of sphericity. Principal Component Analysis (PCA) was then employed to extract initial factors, with the number of factors determined based on eigenvalues greater than one and the scree plot (Alavi et al., 2020). Varimax rotation was applied to simplify the factor structure and enhance interpretability. A minimum factor loading threshold of 0.30 was established to ensure that only items with substantial contributions were retained. Items with cross-loadings where loadings were below 0.30 on multiple factors with a minimal difference between loadings were removed to avoid ambiguity. The process was iterative, involving the removal of problematic items and re-running the PCA to refine the factor structure. After finalizing the factors, they were reviewed for coherence and relevance, ensuring that each item contributed meaningfully to the identified dimensions of cultural bias in health professions.

Reliability. First, the CITC was calculated for each item (Field, 2024). This measure helped determine how well each item correlated with the overall scale score, indicating whether items contributed effectively to the scale's internal consistency. Items with low CITC values were evaluated for potential removal to improve the scale's coherence. Next, Cronbach's alpha was computed to assess the internal consistency of the entire scale. This statistic provided an estimate of how reliably the scale measures the construct it is intended to assess, with a minimum acceptable alpha value of 0.60 set as a benchmark (Kishore et al., 2021). This threshold ensured that the scale had an adequate level of internal consistency. In addition to these measures, split-half reliability testing was performed (DeVellis & Thorpe, 2021). This involved dividing the scale into two halves and comparing the results from each half. By correlating the scores from the two halves, the split-half reliability test evaluated the scale's consistency over different segments, providing further insight into its dependability.

EFA. The factor structure was examined using the EFA. Prerequisites for the analysis included a KMO coefficient of larger than 0.60 and the significance of Bartlett's sphericity test ($p < .05$) (Bartlett, 1954). Varimax rotation and PCA were used to extract common components. Prerequisites were factor loadings of at least 0.30, three items per common factor, and all factors accounting for more than 40% of the variance in the total set (Schreiber, 2021).

CFA. The factor structure identified by the EFA was confirmed by the CFA. Schreiber (2021) provided the following criteria for fit indices: chi-square ratio to degrees of freedom (χ^2/df) $\leq$ 3; root mean square error of approximation

(RMSEA) ≤ 0.05; goodness of fit index (GFI), adjusted goodness of fit index (AGFI), Tucker–Lewis index (TLI), comparative fit index (CFI), and incremental fit index (IFI) > 0.90; parsimonious goodness-of-fit index (PGFI), and parsimonious normed fit index (PNFI) ≥ 0.50. For RMSEA, values below 0.05 indicate good fit, while values below 0.08 represent acceptable fit. For CFI and TLI, values above 0.90 indicate good fit, and values above 0.95 indicate excellent fit. In this study, the fit indices were evaluated by comparing them with commonly accepted thresholds. The χ^2/df value was calculated as 3.178, which slightly exceeds the acceptable threshold of 3, suggesting a marginal fit. The RMSEA value was found to be 0.080, indicating an acceptable fit, although at the upper limit of the acceptable range. The GFI and AGFI values were 0.850 and 0.810, respectively, both falling below the desired threshold of 0.90, indicating a moderate fit. Similarly, the CFI and TLI values were 0.858 and 0.836, respectively, both below the 0.90 threshold, suggesting that the model provides a moderate rather than a good fit. The IFI value of 0.859 also reflects a moderate fit. The lower-than-desired GFI and AGFI values may be attributed to the complexity of the model and the high number of factors, as complex models often result in lower fit indices (Hair & et al, 2010). Despite these lower values, the GFI and AGFI indices are still within an acceptable range considering the model's complexity and the nature of the data. In addition, other fit indices such as RMSEA, CFI, and IFI suggest that the overall model fit is adequate. Future refinements, such as simplifying the model or adjusting the number of factors, could potentially improve the GFI and AGFI indices to achieve a more optimal fit. Future refinements of the model could potentially improve these fit indices, especially the GFI and AGFI, to achieve a more optimal fit. Finally, the PNFI was calculated as 0.700, which is acceptable as it exceeds the minimum threshold of 0.05.

Convergent and Discriminant Validity. Average variance extraction (AVE) and composite reliability (CR) were used to evaluate the instrument's convergent validity. It was necessary for both the AVE and CR values to be higher than 0.50. Subsequently, the discriminant validity was examined by determining if the square root of the AVE for every factor exceeded the correlation coefficients among them (Reichardt & Coleman, 1995).

Internal Consistency and Test–Retest Reliability. Cronbach's α coefficients and split-half reliability coefficients measured internal consistency (Anselmi et al., 2019). To assess test–retest reliability, thirty clinical nurses were selected and reassessed 2 weeks later.

Data Analysis

Data analysis was performed using SPSS version 22 (SPSS Inc.), with a statistical significance level set at $p < .05$. The demographic characteristics of the participants and other key variables were summarized using descriptive statistics (mean, standard deviation, minimum-maximum ranges). The internal consistency of the scale was assessed using Cronbach's alpha coefficient, and the reliability was further evaluated through split-half reliability testing and test–retest reliability analyses (DeVellis & Thorpe, 2021). The performance of individual items was examined using CITC, and differences between the upper and lower 27% of item scores were assessed using independent sample *t*-tests (Yusoff et al., 2021). To determine the factor structure of the scale, PCA with varimax rotation was employed as part of the EFA, and the number of factors was determined based on the Kaiser–Guttman criterion and scree plot (Alavi et al., 2020). CFA was conducted to verify the factor structure obtained from EFA, and the model was evaluated using goodness-of-fit indices (RMSEA, CFI, TLI, and χ^2). Content validity was assessed using the Davis Technique based on expert opinion, and construct validity was established through the results of EFA and CFA (Bennett et al., 2018). Convergent and discriminant validity analyses were performed to test the alignment of the scale with theoretically related and unrelated constructs, while criterion-related validity analyses examined the relationship between the scale and existing validated measures.

Ethical Considerations

Essential approvals were obtained, including ethical clearance from the Scientific Research Ethics Committee (approval code: E.30818). Participants were informed about the study's objectives, potential benefits, and risks, and written informed consent was obtained. All data were anonymized, ensuring confidentiality through secure storage and restricted access. Voluntary participation was emphasized, allowing withdrawal at any time without consequences, and the study was carefully designed to prevent any psychological or physical harm.

Results

Sociodemographic and General Characteristics of Participants

Table 1 shows that of the nurses who took part in the study, 77.6% were female, 55.8% were between the ages of 25 and 35, 50.8% were unmarried, 59.8% had a bachelor's degree, 51.4% had 10–20 years of work experience, and 66.4% had no prior contact with patients from different cultural backgrounds.

Content Validity

Items were revised based on expert opinion to assess the content validity of the scale. Six items with I-CVI values below

Table 1. Frequency Distribution of Demographic Characteristics of Participants.

Characteristics	Participants in the EFA *n*	Participants in the EFA %	Participants in the CFA *n*	Participants in the CFA %
Gender				
Women	249	77.6	242	75.4
Men	72	22.4	79	24.6
Age				
<25	75	23.4	67	20.9
25-35	179	55.8	193	60.1
>35	67	20.9	61	19.0
Marital status				
Unmarried	163	50.8	169	52.6
Married	158	49.2	152	47.4
Educational level				
Junior college	56	17.5	46	14.4
Bachelor's degree	192	59.8	212	66.0
Master degree	73	22.7	63	19.6
Working years				
<10	109	34.0	108	33.6
10-20	165	51,4	172	53.6
<20	47	14.6	41	12.8
Experienced with patient from different cultures				
Yes	108	33.6	113	35.2
No	213	66.4	208	64.8

Note. EFA = exploratory factor analysis; CFA = confirmatory factor analysis.

0.8 (Items 21, 25, 14, 32, 23, 47) were removed from the scale (Davis, 1992), leaving 47 items. Originally, the CITC values were between 0.008 and 0.500. To improve reliability before conducting the EFA, 26 items with CITC values below 0.20 were removed. Table 2 shows the I-CVI scores for the remaining items in the final scale.

Construct Validity

The purpose of the EFA was to look into the scale's underlying factor structure. The KMO score was 0.861, and the object antipatterns were ≥0.5. The results of Bartlett's sphericity test indicated that the data were sufficient for factor analysis (χ^2 = 2,850.732, df = 210, p<.01). The data displayed a normal distribution, and the item skewness and kurtosis values fell within acceptable bounds, making them appropriate for factor analysis. The EFA was performed using PCA with varimax rotation. A three-factor structure for the scale was shown by the eigenvalues, scree plot tests, and total variance explained (Figure 2). The resulting scale had 21 items with factor loadings ranging from 0.341 to 0.862, which together accounted for 49.431% of the variance across the three components. According to Table 2, the percentage of variance explained by each of the separate subscales was 18.582%, 16.342%, and 14.507%.

CFA

The CFA results for the factor model showed the following indices: χ^2/df = 3.178, GFI = 0.850, AGFI = 0.810, RMSEA = 0.080, TLI = 0.836, CFI = 0.858, IFI = 0.859, PNFI = 0.700 (Figure 3). Although the GFI and AGFI values fall below the desired threshold of 0.90, these results suggest that the model is still an adequate fit, considering the complexity and the number of factors involved. The RMSEA value of 0.080 is at the upper limit of the acceptable range, indicating an acceptable fit. The TLI, CFI, and IFI values, although slightly below 0.90, still indicate a moderate fit. The PNFI value of 0.700 exceeds the minimum threshold, which is acceptable. These findings suggest that while the model demonstrates adequate construct validity, future refinements could potentially improve the fit indices, particularly the GFI and AGFI, to achieve a more optimal fit.

Convergent and Discriminant Validity

The analyses of the scale's convergent and discriminant validity supported its hypothesized factor structure. The AVE values ranged from 0.521 to 0.668, while the CR values ranged from 0.882 to 0.981. In addition, the square roots of the AVE values (ranging from 0.49 to 0.81) were higher

Table 2. Psychometric Properties of the Cultural Bias Scale for Health Professionals.

Factors	Items	Cronbach's alpha (α)	I-CVI	*M*	*SD*	CITC total scale	EFA factor loadings
Factor 1. Care Process Factors Cronbach's α = .82 Eigenvalues: 3.902 Explanations variance (%): 18.582	11. Providing care to individuals from different religious, ethnic, racial, and cultural backgrounds makes me unhappy.	.871	0.95	2.44	1.16	0.489	0.469
	12. I do not feel safe when providing care to individuals from different religious, ethnic, racial, and cultural backgrounds	.869	0.92	2.21	1.14	0.533	0.614
	13. I avoid physical contact when providing care to individuals from different religious, ethnic, racial, and cultural backgrounds.	.871	0.93	2.45	1.10	0.464	0.475
	16. I avoid eye contact when providing care to individuals from different religious, ethnic, racial, and cultural backgrounds.	.869	0.90	2.15	1.13	0.542	0.624
	18. I believe that cultural practices performed for the health of individuals from different religious, ethnic, racial, and cultural backgrounds are harmful.	.870	0.90	2.33	1.18	0.499	0.573
	22. I find it difficult to understand the body language of individuals from different religious, ethnic, racial, and cultural backgrounds.	.869	0.96	2.44	1.17	0.531	0.675
	24. I think that individuals from different religious, ethnic, racial, and cultural backgrounds cannot actively participate in their own care.	.871	0.93	2.26	1.13	0.479	0.549
	26. I have difficulty understanding the health perceptions of individuals from different religious, ethnic, racial, and cultural backgrounds.	.870	0.87	2.16	1.29	0.504	0.716
	27. I believe that different cultures negatively impact the quality of health care.	.873	0.86	1.77	0.96	0.395	0.551
	29. I find it difficult to empathize with individuals from different cultures while providing care.	.874	0.91	1.47	0.84	0.392	0.512
	33. I do not want to be tolerant when providing care to individuals from different cultures.	.877	0.92	2.59	1.41	0.336	0.401
Factor 2. Individual Opinion Factors of the Health Professional Cronbach's α = .78 Eigenvalues: 3.902 Explanations variance (%): 16.342	5. I believe my culture is superior to other cultures.	.876	0.94	2.45	1.49	0.376	0.381
	6. I do not want to provide care to someone who criticizes my culture.	.871	0.82	1.41	0.83	0.513	0.746
	7. I realize that I have biases toward the cultures of some individuals while providing care.	.867	0.84	1.69	0.99	0.611	0.789
	8. I do not want to provide care to patients whose culture does not align with or is different from mine.	.866	0.83	1.79	1.06	0.631	0.788
	9. I take a multidimensional approach when providing care to patients from different cultures.	.868	0.87	1.51	0.90	0.618	0.780
	20. I am negatively affected by the social and political events in the country of origin	.874	0.90	1.41	0.88	0.388	0.341
Factor 3. Socioeconomic Factors of the Care Recipient Cronbach's α = .87 Eigenvalues: 3.432 Explanations variance (%): 14.507	34. I provide high-quality health care services to all individuals regardless of their income level.	.877	0.89	1.67	1.10	0.298	0.798
	35. I believe that individuals with a higher income should receive better health care services.	.872	0.89	1.58	1.01	0.45	0.832
	36. I think that individuals receiving social assistance have greater health care needs.	.870	0.87	1.72	1.17	0.513	0.775
	37. I prefer to provide better health care services to individuals who are university graduates or have received education abroad.	.871	0.89	1.50	0.94	0.497	0.862

Note. Explanations variance of total scale: 49.431%. Cronbach's alpha coefficients: α =.858 (first half), α =.755 (second half), and α =.877 (total scale). CITC = corrected item-total correlation; EFA = exploratory factor analysis; I-CVI = the item-level content validity index; *SD* = standard deviation.

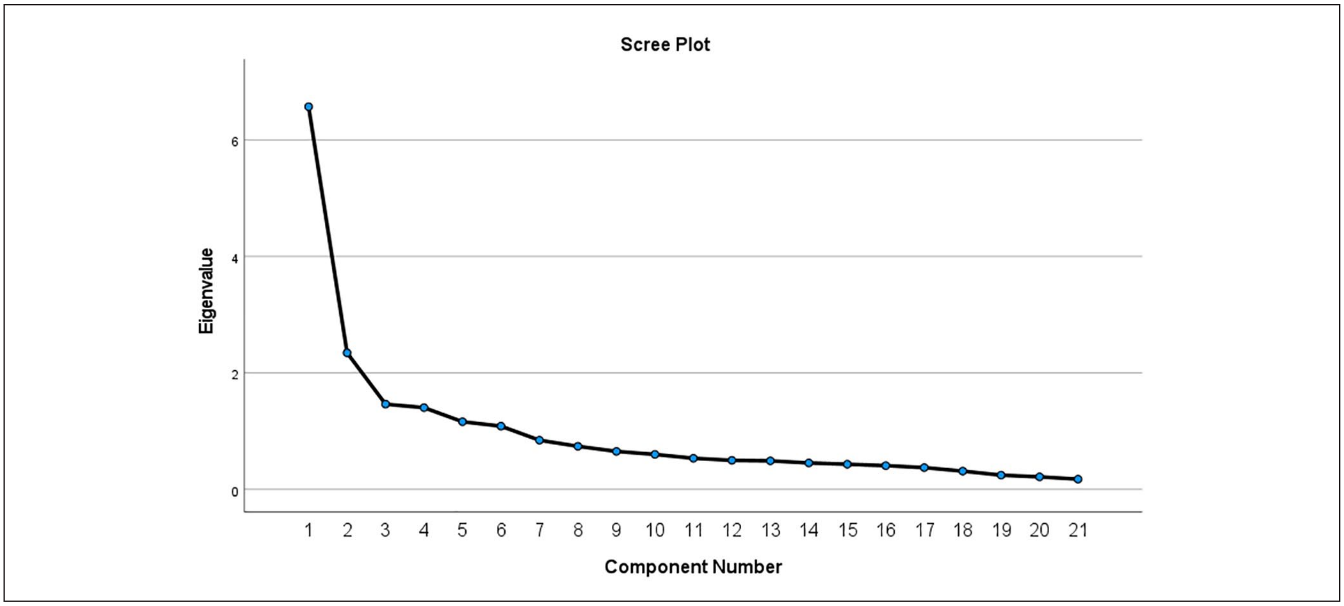

Figure 2. Scree Plot of Exploratory Factor Analysis for the Cultural Bias Scale for Health Professionals.

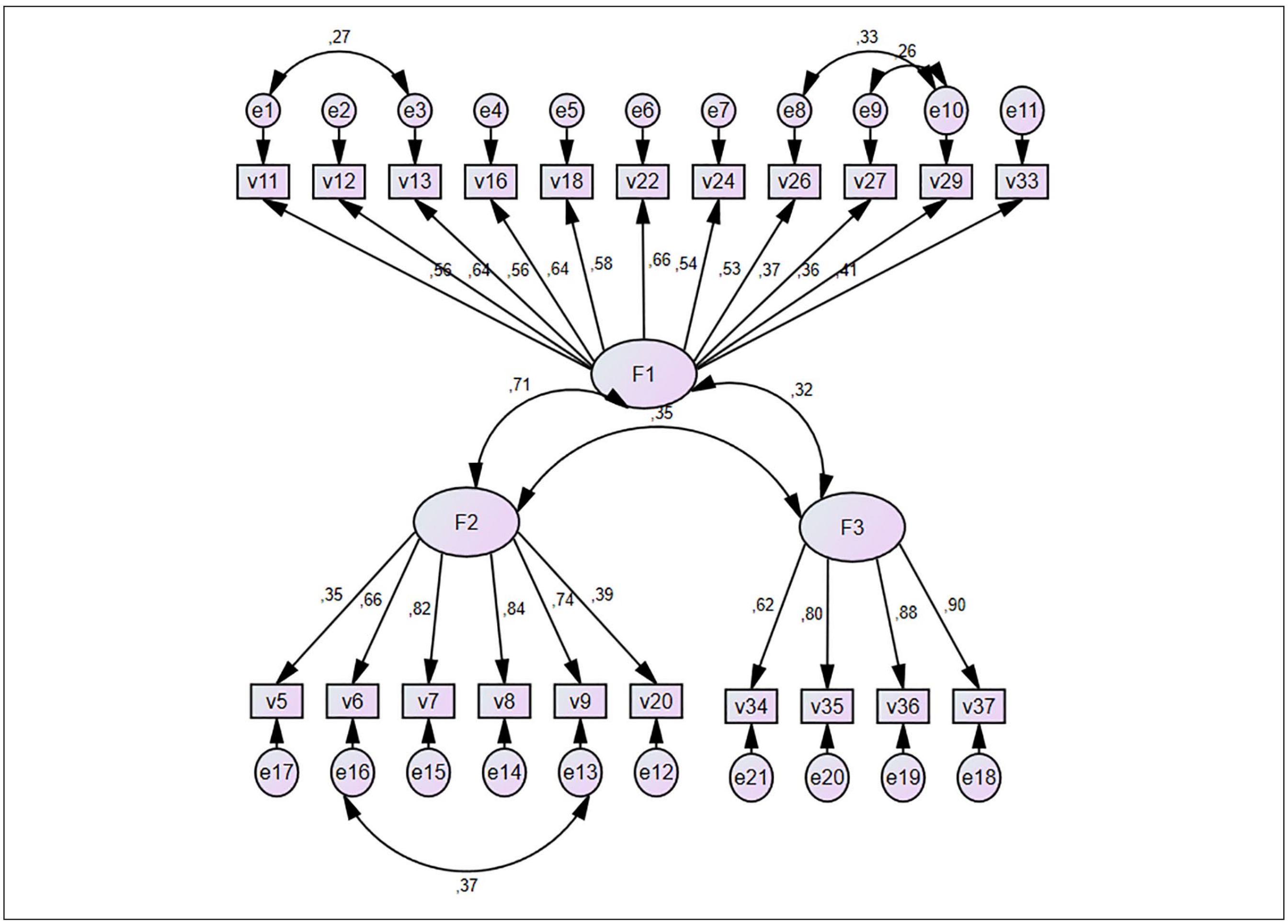

Figure 3. Measurement of the Factor Model of the Cultural Bias Scale for Health Professionals.

Table 3. Convergent Validity and Discriminant Validity of the Cultural Bias Scale for Health Professionals.

	Correlation between factors					
Factors	Factor 1	Factor 2	Factor 3	AVE	Sqtr. AVE	CR
Factor 1	1	0.356	0.288	0.542	0.749	0.963
Factor 2	0.388	1	0.372	0.668	0.866	0.981
Factor 3	0.292	0.372	1	0.521	0.781	0.882

Note. AVE = average variance extracted; CR = construct reliability.

Table 4. Split-Half Reliability of the Cultural Bias Scale for Health.

Cronbach's α	First half	Value	0.808
		N of items	10[a]
	Second half	Value	0.952
		N of items	11[b]
	Total number of items		21
Correlation between two halves			0.791
Spearman–Brown coefficient	Equal length		0.986
	Unequal length		0.896
Guttmann split-half coefficient			0.752

[a]The items are i11, i12, i13, i16, i18, i22, i24, i26, i27, i29. [b] The items are i33, i5, i6, i7, i8, v9, i20, i34, i35, i36, i37.

than the corresponding component correlation coefficients. These findings indicate that the scale has acceptable convergent validity, as the AVE values demonstrate sufficient variance explained for each factor. Furthermore, in terms of discriminant validity, the fact that the square roots of the AVE values were greater than the inter-factor correlations suggests that each factor is sufficiently distinct from the others. This supports the validity of the scale in measuring the intended constructs and clearly differentiating between the factors (Table 3).

Reliability Tests

Internal Consistency (Split-Half Method). The internal consistency of the scale was evaluated using the Cronbach's alpha coefficient, and the reliability of the two halves of the scale was also examined. The internal consistency coefficient for the first half of the scale was found to be 0.808, indicating a good level of reliability. For the second half, a higher coefficient of 0.952 was obtained, reflecting excellent reliability. These results suggest that both parts of the scale have internal consistency, but the second half demonstrates significantly stronger internal consistency. In addition, the correlation coefficient between the two halves of the scale was calculated as 0.791, indicating a moderate to strong relationship between the two parts. The Guttmann split-half coefficient was found to be 0.752, indicating satisfactory split-half reliability. The difference in internal consistency values between the two halves suggests that the items in the second half of the scale may exhibit more homogeneity or stronger cohesion compared to the first half (Table 4).

Discussion

Cultural bias is a critical issue in health care, as cultural differences, language barriers, and biases between health care professionals and patients can hinder effective communication and the delivery of quality care. These barriers can negatively impact processes such as accurate diagnosis, appropriate treatment planning, and patient satisfaction. The challenges nurses face when providing care to immigrant populations and the solutions they propose are critical in addressing health inequalities (Cal et al., 2022). Therefore, measuring the degree of cultural bias can help overcome these inequalities and challenges in health care services. In addition, this scale can be used in services provided to immigrant and minority communities to enhance health care professionals' cultural sensitivity. In the literature, the lack of a measurement tool for culturally intense interactions represents a significant gap. Existing cultural competence scales, such as the Nurse Cultural Competence Scale (NCCS) developed by Perng and Watson (2012), the Cultural Competence Scale for Healthcare Providers developed by Schwarz et al. (2015), and the Cultural Competency Scale for Primary Health Care Professionals (PHCP-CCS) developed by Gozum et al. (2020), have been used to assess the cultural competence of health care professionals (Gozum et al., 2020; Perng & Watson, 2012; Schwarz et al., 2015). However, these scales are limited in measuring cultural bias and intense interactions specifically. This study makes an important contribution by filling this gap. Moreover, this scale provides a novel tool for academic studies aiming to assess cultural biases in health care settings. Similar measures, based on expert opinions and literature reviews, initially resulted in a pool of 53 items. This approach is a comprehensive method used in the development of other scales, such as PHCP-CCS and HPCCI. To enhance the content, readability, and usability of the scale, consultations were held with experts, experienced nurses, nurses participating in the pilot study, and linguists (DeVellis & Thorpe, 2021; Gurbuz & Sahin, 2018; Kishore et al., 2021; Secer, 2018). After 14 experts evaluated the content validity of the scales, it was found that the I-CVI and S-CVI scores were above the acceptable thresholds (>0.8). According to Carpenter (2018), the item pool should initially contain three to four times the final number of items on the scale (Carpenter, 2018). During the EFA phase, as in PHCP-CCS and HPCCI, items with factor loadings below 0.32 were removed, leaving at least three items per subscale. Other researchers, such as Cokluk et al. (2012) and Streiner et al. (2024), recommended removing items with cross-loadings on multiple factors and CITC values <0.2. As a result, all item removal techniques used in this study align with current research. The structural validity of scales is generally tested using factor analysis. Before these analyses, the adequacy of the data and sample size should be evaluated (Cokluk et al., 2012; Streiner et al., 2024). A key measure of sample adequacy is the KMO coefficient, which can be interpreted as follows: 0.7–0.79 (*good*), 0.8–0.9 (*very good*), and >0.9 (*excellent*) (Field, 2024; Secer, 2018). The KMO value obtained in our study was 0.861, indicating a very good and robust sample size, while the KMO value reported in the PHCP-CCS study was 0.84, showing a similar adequacy. In addition, all item skewness and kurtosis values were within acceptable limits, demonstrating normal distribution, and Bartlett's test was found to be significant ($p < .01$) (Gurbuz & Sahin, 2018; Kline, 2023).

In our study, the structural validity of the scale developed to measure cultural bias was tested through exploratory and confirmatory factor analyses. The results showed that the factor structure of the scale was robust and explained a significant portion of the total variance. The focus on measuring cultural bias specifically may account for the differences found when compared to previous cultural competence scales. Moreover, the CFA results showed that the model fit was within acceptable limits in the literature. These findings confirm that the scale we developed is both valid and reliable (Alavi et al., 2020; Carpenter, 2018; Cokluk et al., 2012; Secer, 2018). The analyses of the scale's convergent and discriminant validity confirmed the assumed factor structure. The range of AVE values was between 0.521 and 0.668, while the range of CR values was between 0.882 and 0.981. These values show that our scale is strong in terms of convergent and discriminant validity, comparable to or better than other scales. The fact that the square root of the AVE values (ranging from 0.49 to 0.81) exceeded the correlation coefficients among similar factors demonstrated good convergent and discriminant validity (Reichardt & Coleman, 1995). Consistent with other research, factor loadings ranged from 0.341 to 0.862, explained more than 40% of the total variance, and each component contributed at least 5% (Cokluk et al., 2012; Secer, 2018). Cronbach's alpha coefficient is commonly used to evaluate the reliability of Likert-type scales (Kilic, 2016). For new scales, $\alpha > .6$ is acceptable, but $\alpha > .7$ is a desired value (Kishore et al., 2021). The α-values for our scale, including for both the first and second halves, were above .7, demonstrating strong internal consistency. These values are comparable to the $\alpha = .84$ for the total scale and α values ranging from .76 to .87 for the subscales of the PHCP-CCS. Split-half tests provided additional evidence for the reliability of both scales. These findings show that our scale is equal to or better than other cultural competence scales in terms of reliability. Thus, the scale demonstrated good internal consistency and homogeneity according to the literature (DeVellis & Thorpe, 2021; Kilic, 2016; Kishore et al., 2021).

Practice Recommendations

The developed scale can be used in various health care settings, including primary care, acute care, and long-term care, to assess health care professionals' cultural biases and improve patient-centered care. It can help evaluate the impact of cultural biases on patient satisfaction and treatment adherence, making it valuable for quality improvement initiatives. In addition, integrating the scale into institutional policies can enhance cultural sensitivity and reduce health care disparities, particularly in services for immigrant and minority communities. Health care administrators can also use it for periodic assessments, ensuring continuous improvement in cultural competence. Incorporating the scale into professional development programs can help health care professionals recognize and manage their cultural biases, enhancing patient-centered care. Nursing and health care education institutions can use it to assess students' cultural sensitivity and refine curricula accordingly. In addition, the scale provides a reliable way to evaluate the effectiveness of cultural competency training, supporting data-driven improvements in education and practice. Regular use by administrators and training departments can help build a more culturally competent workforce, ultimately improving health care outcomes.

Limitations

A significant limitation of this study is that the sample consists solely of nurses, limiting the generalizability of the findings to other health care professionals who may experience cultural bias differently. In addition, the scale was developed within the Turkish context, specifically in Erzurum, a region with a unique socio-cultural structure that may influence health care professionals' biases. As a result, findings may vary in different geographical regions or hospitals, reducing their applicability to diverse health care settings. Another limitation is the voluntary participation of health care professionals, which may lead to sampling bias by including individuals with higher cultural bias awareness. To enhance the validity and reliability of the scale, future studies should involve more diverse health care professionals and be conducted across multiple centers with different cultural contexts, improving the external validity and applicability of the findings.

Conclusion

The developed scale consists of 21 items across three subscales, assessing the challenges health care professionals face in providing care to diverse cultural, ethnic, and socioeconomic groups. The Care Process Factors subscale (11 items) evaluates the impact of cultural barriers on care quality. The Individual Opinion Factors of the Health Professional subscale (6 items) measures how personal biases and cultural sensitivity influence patient care. The Socioeconomic Factors of the Care Recipient subscale (4 items) examines how patients' socioeconomic status affects health care access and quality. The total score ranges from 21 to 105, with higher scores indicating greater cultural prejudice. As a valid and reliable tool, this scale supports efforts to enhance cultural sensitivity in health care services.

Author Contributions

E.K., A.Y., T.K.S., and Ü.S. contributed to the writing of the main manuscript. E.K., T.K.S., and Ü.S. collected the data. E.K. analyzed the data. A.Y. supervised the study. All authors reviewed the manuscript, provided thorough feedback, and approved the final version.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Esin Kavuran https://orcid.org/0000-0002-4890-520X

Afife Yurttaş https://orcid.org/0009-0000-0557-7392

References

Alavi, M., Visentin, D. C., Thapa, D. K., Hunt, G. E., Watson, R., & Cleary, M. (2020). Chi-square for model fit in confirmatory factor analysis. *Journal of Advanced Nursing*, *76*(9), 2209–2211. https://doi.org/10.1111/jan.14399

Almanasreh, E., Moles, R., & Chen, T. F. (2019). Evaluation of methods used for estimating content validity. *Research in Social and Administrative Pharmacy*, *15*(2), 214–221. https://doi.org/10.1016/j.sapharm.2018.03.066

Almutairi, A. F., Adlan, A. A., & Nasim, M. (2017). Perceptions of the critical cultural competence of registered nurses in Canada. *BMC Nursing*, *16*, 1–9. https://doi.org/10.1186/s12912-017-0242-2

Anselmi, P., Colledani, D., & Robusto, E. (2019). A comparison of classical and modern measures of internal consistency. *Frontiers in Psychology*, *10*, 492664. https://doi.org/10.3389/fpsyg.2019.02714

Bartlett, M. S. (1954). A note on the multiplying factors for various χ2 approximations. *Journal of the Royal Statistical Society. Series B (Methodological)*, *16*, 296–298.

Baumgarten, E. (2018). From Kurdistan to Baghdad: The transfer of visual knowledge during the early modern period. *Ars Judaica: The Bar-Ilan Journal of Jewish Art*, *14*(2018), 79–92.

Bennett, A. A., Bakker, A. B., & Field, J. G. (2018). Recovery from work-related effort: A meta-analysis. *Journal of Organizational Behavior*, *39*(3), 262–275. https://doi.org/10.1002/job.2217

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

Bolarinwa, O. A. (2015). Principles and methods of validity and reliability testing of questionnaires used in social and health science researches. *Nigerian Postgraduate Medical Journal*, *22*(4), 195–201. https://doi.org/10.4103/1117-1936.173959

Cal, A., Cavusoglu, F., & Aydın Avci, İ. (2022). Problems experienced by nurses while providing care to immigrants and suggestions for solutions Hemşirelerin Göçmen Bireylere Bakım Verirken Yaşadıkları Sorunlar ve Çözüme Yönelik Öneriler. *E-Journal of Dokuz Eylul University Nursing Faculty*, *15*(2), 197–206. https://doi.org/10.46483/deuhfed.804655

Campinha-Bacote, J. (2002). The process of cultural competence in the delivery of healthcare services: A model of care. *Journal of Transcultural Nursing*, *13*(3), 181–184. https://doi.org/10.1177/10459602013003003

Carpenter, S. (2018). Ten steps in scale development and reporting: A guide for researchers. *Communication Methods and Measures*, *12*(1), 25–44. https://doi.org/10.1080/19312458.2017.1396583

Cokluk, O., Sekercioglu, G., & Buyukozturk, S. (2012). *Multivariate statistics for social sciences: SPSS and LISREL applications* (Vol. 2). Pegem Academy.

Davis, L. L. (1992). Instrument review: Getting the most from a panel of experts. *Applied Nursing Research*, *5*(4), 194–197. https://doi.org/10.1016/S0897-1897(05)80008-4

De Haas, H., Castles, S., & Miller, M. J. (2019). *The age of migration: International population movements in the modern world*. Bloomsbury Publishing.

DeVellis, R. F., & Thorpe, C. T. (2021). *Scale development: Theory and applications*. SAGE.

Duffy, J. R. (2022). *Quality caring in nursing and health systems: Implications for clinicians, educators, and leaders*. Springer Publishing Company.

Durand, F., Bourgeault, I. L., Hebert, R. L., & Fleury, M.-J. (2022). The role of gender, profession and informational role self-efficacy in physician–nurse knowledge sharing and decision-making. *Journal of Interprofessional Care*, *36*(1), 34–43. https://doi.org/10.1080/13561820.2021.1890006

Field, A. (2024). *Discovering statistics using IBM SPSS Statistics*. SAGE.

Gozum, S., Tuzcu, A., & Yurt, S. (2020). Developing a cultural competency scale for primary health care professionals. *Studies in Psychology*, *40*(2), 429–450. https://doi.org/10.26650/SP2019-0043

Gurbuz, S., & Sahin, F. (2018). *Research methods in social sciences*. Seckin Hukuk.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Pearson Education.

Karatas, A. (2023). Global migration governance: An analysis of turkey on the basis of migration governance indicators. In İ. Dursunoğlu (Ed.), *International research in social, human and administrative sciences X* (pp. 47–90). Eğitim Yayınevi.

Kaya, A., & Boz, İ. (2019). The development of the professional values model in nursing. *Nursing Ethics*, *26*(3), 914–923. https://doi.org/10.1177/0969733017730685

Kilic, S. (2016). Cronbach's alpha reliability coefficient. *Psychiatry and Behavioral Sciences*, *6*(1), 47–48. https://doi.org/10.5455/jmood.20160307122823

Kirmayer, L. J., & Jarvis, G. E. (2019). Culturally responsive services as a path to equity in mental healthcare. *HealthcarePapers*, *18*(2), 11–23. https://doi.org/10.12927/hcpap.2019.25925

Kishore, K., Jaswal, V., Kulkarni, V., & De, D. (2021). Practical guidelines to develop and evaluate a questionnaire. *Indian Dermatology Online Journal*, *12*(2), 266–275. https://doi.org/10.4103/idoj.IDOJ_674_20

Kline, R. B. (2023). *Principles and practice of structural equation modeling*. Guilford Publications.

Kressin, N. R., Raymond, K. L., & Manze, M. (2008). Perceptions of race/ethnicity-based discrimination: A review of measures and evaluation of their usefulness for the health care setting. *Journal of Health Care for the Poor and Underserved*, *19*(3), 697–730. https://doi.org/10.1353/hpu.0.0041

Kyriazos, T. A., & Stalikas, A. (2018). Applied psychometrics: The steps of scale development and standardization process. *Psychology*, *9*(11), 2531–2560. https://doi.org/10.4236/psych.2018.911145

Leyva-Moral, J., & Bernabeu-Tamayo, M. D. (2021). The theoretical background and history of transcultural nursing. In A. Yava & B. Tosun (Eds.), *Transcultural nursing benefits: Better and effective nursing education for improving transcultural nursing skills* (1st ed., pp. 32–51). Ankara Nobel Tıp Kitabevleri.

Lima, A. F. S., Santos, C. E. B., Alves, N. R., Lima, M. C. F., Jorge, J. S., Tigre, H. W. A., de Almeida, A. V. A., da Silva Santos, T., & de Miranda Crispim Costa, L. (2023). Nursing care for the Warao people: An experience report based on transcultural theory. *Revista da Escola de Enfermagem da USP*, *57*, Article e20230035.

Perng, S. J., & Watson, R. (2012). Construct validation of the nurse cultural competence scale: A hierarchy of abilities. *Journal of Clinical Nursing*, *21*(11–12), 1678–1684. https://doi.org/10.1111/j.1365-2702.2011.03933.x

Polit-O'Hara, D., & Yang, F. M. (2016). *Measurement and the measurement of change: A primer for the health professions*. Wolters Kluwer.

Reichardt, C. S., & Coleman, S. (1995). The criteria for convergent and discriminant validity in a multitrait-multimethod matrix. *Multivariate Behavioral Research*, *30*(4), 513–538. https://doi.org/10.1207/s15327906mbr3004_3

Rew, L., Becker, H., Cookston, J., Khosropour, S., & Martinez, S. (2003). Measuring cultural awareness in nursing students. *The Journal of Nursing Education*, *42*, 249–257.

Sapnas, K. G., & Zeller, R. A. (2002). Minimizing sample size when using exploratory factor analysis for measurement. *Journal of Nursing Measurement*, *10*(2), 135–154. https://doi.org/10.1891/jnum.10.2.135.52552

Sastrawan, S., Newton, J. M., & Malik, G. (2019). Nurses' integrity and coping strategies: An integrative review. *Journal of Clinical Nursing*, *28*(5–6), 733–744. https://doi.org/10.1111/jocn.14702

Schreiber, J. B. (2021). Issues and recommendations for exploratory factor analysis and principal component analysis. *Research in Social and Administrative Pharmacy*, *17*(5), 1004–1011. https://doi.org/10.1016/j.sapharm.2020.07.027

Schumacker, R. E., & Lomax, R. G. (2010). *A beginner's guide to structural equation modeling* (3rd ed.). Routledge.

Schwarz, J. L., Witte, R., Sellers, S. L., Luzadis, R. A., Weiner, J. L., Domingo-Snyder, E., & Page Jr, J. E. (2015). Development and psychometric assessment of the healthcare provider cultural competence instrument. *INQUIRY: The Journal of Health Care Organization, Provision, and Financing*, *52*, 0046958015583696. https://doi.org/10.1177/0046958015583696

Secer, I. (2018). *Psychological test development and adaptation process: SPSS and LISREL applications*. ANI Publications.

Streiner, D. L., Norman, G. R., & Cairney, J. (2024). *Health measurement scales: A practical guide to their development and use*. Oxford University Press.

Tang, C., Tian, B., Zhang, X., Zhang, K., Xiao, X., Simoni, J. M., & Wang, H. (2019). The influence of cultural competence of nurses on patient satisfaction and the mediating effect of patient trust. *Journal of Advanced Nursing*, *75*(4), 749–759. http://doi.org/10.1111/jan.13854

Tisdale, D., & Symenuk, P. M. (2020). Human rights and nursing codes of ethics in Canada 1953–2017. *Nursing Ethics*, *27*(4), 1077–1088. https://doi.org/10.1177/0969733020906606

Yusoff, N. S., Kaman, Z. K., Zahari, A. R., Norafi, W. H. W. M., & Abdullah, A. B. (2021). Examining smart meter users' experience on continuance intention in adopting smart meter in Malaysia—Result from a pilot study. *Asia Proceedings of Social Sciences*, *7*(2), 110–113. https://doi.org/10.31580/apss.v7i2.1785