

Reliability and Validity of the Turkish Version of the Job Performance Scale Instrument

Arzu Kader Harmanci Seren¹ • Rujnan Tuna^{2*} • Feride Eskin Bacaksiz³

ABSTRACT

Background: Objective measurement of the job performance of nursing staff using valid and reliable instruments is important in the evaluation of healthcare quality. A current, valid, and reliable instrument that specifically measures the performance of nurses is required for this purpose.

Purpose: The aim of this study was to determine the validity and reliability of the Turkish version of the Job Performance Instrument.

Methods: This study used a methodological design and a sample of 240 nurses working at different units in four hospitals in Istanbul, Turkey. A descriptive data form, the Job Performance Scale, and the Employee Performance Scale were used to collect data. Data were analyzed using IBM SPSS Statistics Version 21.0 and LISREL Version 8.51.

Results: On the basis of the data analysis, the instrument was revised. Some items were deleted, and subscales were combined.

Conclusions/Implications for Practice: The Turkish version of the Job Performance Instrument was determined to be valid and reliable to measure the performance of nurses. The instrument is suitable for evaluating current nursing roles.

KEY WORDS:

contextual performance, job performance, nurses, reliability and validity, task performance.

The job performance of nurses is a multidimensional concept (Bakker, Demerouti, & Euwema, 2005; Coleman & Borman, 2000; Griffin, Neal, & Parker, 2007). In academic studies, nurse performance is usually approached as two components, namely, task performance and contextual performance (Bakker et al., 2005; Borman & Motowidlo, 1993; Coleman & Borman, 2000). Task performance includes the roles of employees within an organization in terms of their main jobs and tasks (Coleman & Borman, 2000) and activities that support the main functions of healthcare institutions (mostly hospitals) and contribute to the achievement of primary targets, with their content usually revealed through business analyses (Borman & Motowidlo, 1993). The second component, contextual performance, is stated as an employee's willingness to help colleagues, ability to achieve collaboratively, and willingness to exert extra effort to complete the work (Coleman & Borman, 2000).

Although nurse performance has a significant impact on the delivery of qualified healthcare services, only a few instruments have been developed for use in this field. The most commonly known are the "Schwirian Six-D Scale" (Schwirian, 1978) and the "Slater Nursing Competencies Rating Scale" (Wandelt & Stewart Slater, 1975). Both were developed in the 1960s and 1970s (Redfern & Norman, 1990; Wandelt & Phaneuf, 1972). These instruments have largely met the need to evaluate nurse performance for many years. However, the role of nurses within the health system today has expanded, the delivery of qualified healthcare services continues to grow in importance, and new concepts such as the satisfaction with and expectations of patients and healthy individuals regarding healthcare services have come into play. Evaluation of the performance levels of nurses who play main roles in the delivery of healthcare services is of critical importance. In light of all these developments and changes, new methods and instruments to evaluate the job performance of nurses are required (Hamilton et al., 2007; Pelletier et al., 2000). Although healthcare institutions have their own performance scales to conduct this evaluation, there is no valid and reliable performance scale for measuring the performance of nurses in Turkey.

Introduction

Nursing care affects the quality of healthcare that is provided to patients in the acute care setting. Technical knowledge and skills, experience, educational level, and efficiency of service provision, that is, nursing "performance," play a very important role in achieving desired healthcare outcomes (Whyte, Lugton, & Fawcett, 2000).

Healthcare services are complex and provided by a multidisciplinary team. Nurses play critical roles within a patient care team in terms of both their number and the services that they provide. Nurses spend most of their time with patients and their relatives and affect health outcomes directly through nursing care (DeLucia, Ott, & Palmieri, 2009; Larrabee et al., 2004; Pappas, 2008). To realize this positive effect and achieve expected outcomes, nurses should perform competently, and managers should evaluate their performance (Behrenbeck, Timm, Griebenow, & Demmer, 2005; Gregg, 2002).

¹PhD, RN, Assistant Professor, Florence Nightingale Faculty of Nursing, Department of Nursing Administration, Istanbul University, Istanbul, Turkey • ²PhD, RN, Research Assistant, Faculty of Health Sciences, Department of Nursing, Istanbul Medeniyet University, Istanbul, Turkey • ³PhD, RN, Research Assistant, Florence Nightingale Faculty of Nursing, Department of Nursing, Administration, Istanbul University, Istanbul, Turkey.

In this study, all of the abovementioned scales that had been developed for the purpose of measuring nurses' performance were evaluated in terms of their validity and reliability. The Turkish adaptation of the scale that was developed by Greenslade and Jimmieson (2007) was preferred, as it was developed after 2000 and seemed to more effectively address the concepts currently used in the field of nursing care.

Methods

Design

A cross-sectional and methodological study design was used for the purposes of adapting the Job Performance Scale (JPS) into Turkish and evaluating its psychometric properties.

Sample

The sample consisted of 240 volunteers, all of whom were nurses working at four hospitals (a public hospital, a public university hospital, a private hospital, and a private university hospital) in Istanbul, Turkey.

Data Collection

Data were collected at the hospitals between March and June 2015. A descriptive data collection form, the JPS, and the Employee Performance Scale (EPS), which was used to show criterion validity, were used for data collection.

Descriptive data collection form

This form consists of seven questions regarding the respondent's age, gender, educational level, unit position, and tenure both in the hospital and the nursing profession.

Job Performance Scale

The original version of the JPS, developed by Greenslade and Jimmieson (2007), was based on the job performance model of Borman and Motowidlo (1993) and intended to measure the performance of nurses. Greenslade and Jimmieson aimed to develop a performance scale to be used on nurses and based it on a structure that approached the job performance concept of Borman and Motowidlo as two different domains: task performance and contextual performance. In the JPS, task performance is an aspect of performance that directly contributes to the organization's technical competence and is related to the employee's work, whereas contextual performance is examined as a performance aspect that expands the organization's social environment and impact and includes the employee's voluntary behaviors (Borman & Motowidlo, 1993). The original scale for which the authors determined the items according to focus group interviews consists of two subscales. The first of these subscales is the Task Performance Scale (TPS), and the second is the Contextual Performance Scale (CPS). The TPS consists of the four subscales

of information (seven items), coordination of care (five items), social support (six items), and technical care (five items), respectively, with a total of 23 items. The Cronbach's alpha internal consistency coefficients of this scale and its subscales vary between .85 and .94. The CPS consists of the four subscales of interpersonal support (six items), job task support (six items), compliance (three items), and volunteering for additional duties (four items), respectively, with a total of 18 items. The Cronbach's alpha internal consistency coefficients of the scale and its subscales vary between .80 and .90. For this scale, which was developed as a 7-point Likert-type scale, each statement is scored between 1 and 7 points, with a higher item total mean score associated with better performance on that item (Greenslade & Jimmieson, 2007).

Employee Performance Scale

The EPS, which is widely used to assess criterion validity, was developed by Erdoğan (2011). The employee performance subscale consists of seven items, and the Cronbach's alpha for the EPS was .94 in the current study. Statements in the 5-point Likert-type scale are scored on a spectrum between "absolutely disagree" (1 point) and "absolutely agree" (5 points), with higher total mean scores for an item associated with better performance on that item.

Procedures

No universal agreement exists regarding adaptation of an instrument for use in a different cultural setting (Gjersing, Caplehorn, & Clausen, 2010). Different numbers of similar steps in terms of context have been suggested for scale adaptation studies (Borsa, Damásio, & Bandeira, 2012; Gjersing et al., 2010). This study adopted the steps that were suggested by Gjersing et al. (2010) for scale adaptation studies (Table 1).

Data Analysis

The obtained data were analyzed using IBM SPSS Statistics Version 21 (IBM Inc., Armonk, NY, USA) and LISREL Version 8.51 (Scientific Software International, Skokie, IL, USA). Descriptive statistics (number, percentage, mean, standard deviation), correlations (Pearson product-moment correlation), and psychometric tests (content validity ratio, item-total correlation, Kaiser-Meyer-Olkin [KMO] measure of adequacy and Bartlett's test of sphericity, exploratory factor analysis [EFA], and confirmatory factor analysis [CFA], internal consistency coefficient) were used for the data analysis.

Ethical Consideration

The permission of the researchers who developed the original scale was received via e-mail. Ethics committee approval

TABLE 1.***The Study Procedures According to “A Suggested Cross-Cultural Adaptations Process” by Gjersing et al. (2010)***

Suggested	Performed
Adaptation process	
Investigation of conceptual and item equivalence	Literature review Discussion with experts in the field and members of target population
Original instrument translated	Two independent translators Fluent in Turkish and good understanding of English
A synthesized translated version	One academician Good command of both languages
Back-translations	Two independent translators Fluent in Turkish and good understanding of English
A synthesized back-translated version	One academician Good command of both languages
Expert committee	12 experts In the area of nursing and had experiences of instrument development and adaptation
Instrument pretested	17 voluntaries Same characteristics with target population
Revised instrument	Researchers
Investigation of operational equivalence	Discussions with experts in the field of nursing
Validation process	
Main study	240 voluntary nurses
Exploratory and confirmatory factor analysis	CFA for testing the fit of the original structure on target population Item analyses regarding the reliability of each item EFA for exploring the factor structure on target population CFA for confirming the construct validity Cronbach’s alpha for testing the internal consistency Correlation analyses for criterion validity
Final instrument	

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

was received with the Decision no. A.07 dated January 6, 2015, from the ethics committee of a university in Istanbul. Approvals were also received from the administrative and nursing service departments of the hospitals where the data were collected.

Results

Participants

The participants were mostly female (94.6%), held a bachelor’s degree (62.1%), and worked in inpatient units (79.1%) as bedside nurses (85.4%). The ages of participants varied from 19 to 52 years (mean = 31.95 years, *SD* = 7.55 years), and their experience at the hospital and in the nursing profession ranged from 1 to 33 years (mean = 8.11 years, *SD* = 7.92 years) and 1 to 33 years (mean = 10.29 years, *SD* = 8.14 years), respectively.

Adaptation Process

Language validity was conducted for both TPS and CPS. This stage was based on the method recommended by the World Health Organization for the adaptation of instruments that were developed in a language different from the target language (World Health Organization, 2008). The original scale was translated and back-translated by four language professionals. Translated and back-translated versions of the scales were then synthesized by two academicians who had a good command of both Turkish and English. Finally, the original scale and the translated version were compared by an expert committee for content validity. The Lawshe technique was used to assess content validity (Lawshe, 1975). The prepared Turkish form was evaluated by 12 experts outside the research team who both worked in the field of nursing and had experience with scale development or adaptation studies. As opinions were received from 12 experts, as suggested by the Lawshe technique, the content validity criterion

was specified as .56 (Lawshe, 1975). No items were omitted at this stage, as no statement received a value below .56.

A pilot study was conducted using a 17-person group outside the sample group in February 2015. Respondents were probed for their understanding to identify potentially confusing or misleading items. On the basis of their responses, the items were finalized by the researchers and discussed with four nurse managers in terms of performance evaluation.

Validation Process

The final version of the adapted scale was given to the main sample of 240 people.

Construct Validity

The CFA was conducted for both scales primarily to evaluate construct validity. The CFA did not confirm the factor structures of the original scales. The realization of the recommended modifications did not ensure any increase at an acceptable level in the goodness-of-fit indices; whereupon, an additional CFA that approached both scales as a single subscale was conducted as an alternative. However, the results indicated that the goodness-of-fit indices did not improve but rather fell to lower levels (Table 2).

Item Analysis

When the item–total score correlations for the total of 41 items, including 23 in the TPS and 18 in the CPS, were examined, the item–total score correlation coefficients for Item 7 in the TPS and Items 2, 8, 11, 16, and 17 in the CPS were found to be $r < .45$. For this reason, these six items were omitted.

A new EFA was then conducted that considered all of the items of both scales as a single item pool (principal component analysis/varimax rotation). At this stage, a KMO test and Bartlett’s test of sphericity were performed, finding a KMO value of .89 and a sphericity test result of $\chi^2 = 7206.1$

($df = 595, p < .001$). The difference between the factor load values was less than .10. Thus, as the EFA results showed that three items loaded in two or more factors simultaneously, these items were also omitted from the scale. The factor loads of the 32 items remaining in the scale varied between .50 and .84, and the scale items were distributed to six factors having eigenvalues above 1 and explaining 70.6% of the total variance. Nine items were involved in the first factor, seven items were involved in the second factor, six items were involved in the third factor, four items each were involved in the fourth and fifth factors, and two items were involved in the sixth factor (Table 3).

CFA was applied again to evaluate the fit of the newly explored structure of the scale. When modification suggestions were examined, error covariance was assigned between Items 8–9 and 22–23 in the first factor, Items 13–14 and 16–18 in the second factor, and Items 24–27 in the third factor (Figure 1). Factor loadings in the subscales were found to be $\geq .64$ in the first factor, $\geq .66$ in the second factor, $\geq .59$ in the third factor, $\geq .56$ in the fourth factor, $\geq .36$ in the fifth factor, and $\geq .65$ in the sixth factor (Figure 1). Goodness-of-fit indices were calculated as $\chi^2 = 1828.22$, $df = 453$, root mean square error of approximation (RMSEA) = .09, goodness-of-fit index (GFI) = .91, adjusted GFI (AGFI) = .88, and comparative fit index (CFI) = .95 (Table 2).

Internal Consistency Analysis

The Cronbach’s alpha reliability coefficients of the subscales that were obtained after the factor analysis and the subscales in the analysis that were conducted to evaluate the scale’s total internal consistency varied between .65 and .93. The total score for the scale was $\alpha = .95$ (Table 3).

Criterion Validity

The EPS was applied on the same group simultaneously to determine the criterion validity. After testing for linearity, Pearson correlation analysis was applied, and positive (+), moderate level ($r = .617$), and advanced level significant

TABLE 2.
Goodness of Fit Indices for the Original and Adapted Version of the Scales

Scale	Factor loadings	χ^2	df	RMSEA	GFI	AGFI	CFI
Original Scales							
Task Performance Scale							
4-Factor	$\geq .66$	1993.60	226	.181	.58	.49	.71
One Factor	$\geq .60$	4576.75	230	.281	.38	.25	.51
Contextual Performance Scale							
4-Factor	$\geq .59$	1794.30	131	.230	.55	.41	.65
One Factor	$\geq .53$	3003.42	135	.298	.42	.26	.44
Adapted Version of the Scale							
Nurse Job Performance Scale							
		1828.22	453	.094	.91	.88	.95

Note. *df* = degrees of freedom; RMSEA = root mean square error of approximation; GFI = goodness of fit index; AGFI = adjusted goodness of fit index; CFI = comparative fit index.

TABLE 3.
Results of Reliability and Structural Analyses

Factor Number	Item	n	% of Explained Variance	Factor Loading	Corrected Item Total Correlations	Cronbach's Alpha
F1	8, 9, 10, 11, 12, 19, 21, 22, 23	9	38.90	.562–.844	.66–.82	.93
F2	13, 14, 15, 16, 17, 18, 20	7	11.13	.593–.840	.63–.85	.93
F3	24, 25, 26, 27, 28, 29	6	8.04	.701–.836	.63–.80	.88
F4	30, 37, 38, 41	4	4.83	.631–.805	.58–.77	.82
F5	1, 2, 4, 6	4	4.17	.593–.649	.51–.73	.79
F6	33, 35	2	3.54	.504–.610	.50–.50	.65
Total		32	70.59	.504–.844	.45–.71	.95

($p < .001$) correlations were determined between the total scores obtained from both scales.

Discussion

Adaptation Process

Language validity

The scale adaptation studies referenced in the literature recommend implementing practices to diminish the psycholinguistic differences between cultures (Şencan, 2005). For this purpose, the translations of the scale first from the original English into Turkish and then from Turkish back into English were conducted in this study in accordance with the method recommended by the World Health Organization (2008) for rendering instruments into languages other than the original language.

Content validity

In scale development and adaptation studies, content validity is used to assess whether the items of a scale include the field to be measured (Öner, 2008). The Turkish form prepared within this study was submitted for the opinion of experts. The content validity criterion was found to be .56 (Lawshe, 1975), and no items were omitted in this stage, as no statement had a criterion value below this value.

Validation Process

First stage: confirmatory factor analysis

Factor analysis is the most commonly used method of analysis to test the construct validity of a scale. Factor analysis is a concept that relates to how a scale accurately measures what (Öner, 2008). CFA was conducted in this study to test the structure of the scale.

In the CFA, the various results of the GFI determine the compatibility of the model. There are numerous goodness-of-fit indices, and there is no absolute consensus as to which ones should be reported (Şimşek, 2007). Chi-square/degree of freedom, RMSEA, AGFI, GFI, and CFI results

were reported in this study as the most commonly used fit indices. According to the results of the analyses, it was determined that adaptation values did not show acceptable fit (Table 2).

Second stage: item analysis

Item analysis gives information regarding the reliability of each item in a scale. Therefore, if items in a scale are equally weighted and independent, the correlation value between each item and the total score should be high (Tavşancıl, 2014). The item–total score correlations for the items in the scale's Turkish form were evaluated as part of the item analysis that was conducted for this study. Accordingly, as the correlation values for the total of 41 items in both scales were $r < .45$, six items were omitted from the study.

Third stage: exploratory factor analysis

As the CFAs did not confirm the original structure, a new EFA was carried out that considered all items of both scales, which had been reduced to 35 items after the item analysis, as a single item pool.

The literature recommends applying a KMO test and a Bartlett's test of sphericity before performing factor analysis. In this study, the KMO coefficient was evaluated as "good," and the Bartlett's test result was determined to be significant in the advanced level (Şencan, 2005).

The literature further recommends that, if each item is listed under a factor, where it receives the highest factor load value, after the EFA; however, the items receiving high factor loads are evaluated under multiple factors; and the difference between the factor load values obtained by the same item under different factors is lower than 0.10, the item in question should be omitted (Büyüköztürk, 2011). Item 5 in the TPS and Items 9 and 13 in the CPS of the original form were omitted in this study because they received similar factor load values in two or more factors. The factor load values of the remaining 32 items of the scale varied between .50 and .84 (Table 3), and the scale items were distributed across six factors, which explained 70.6% of the total variance. The items in the "coordination

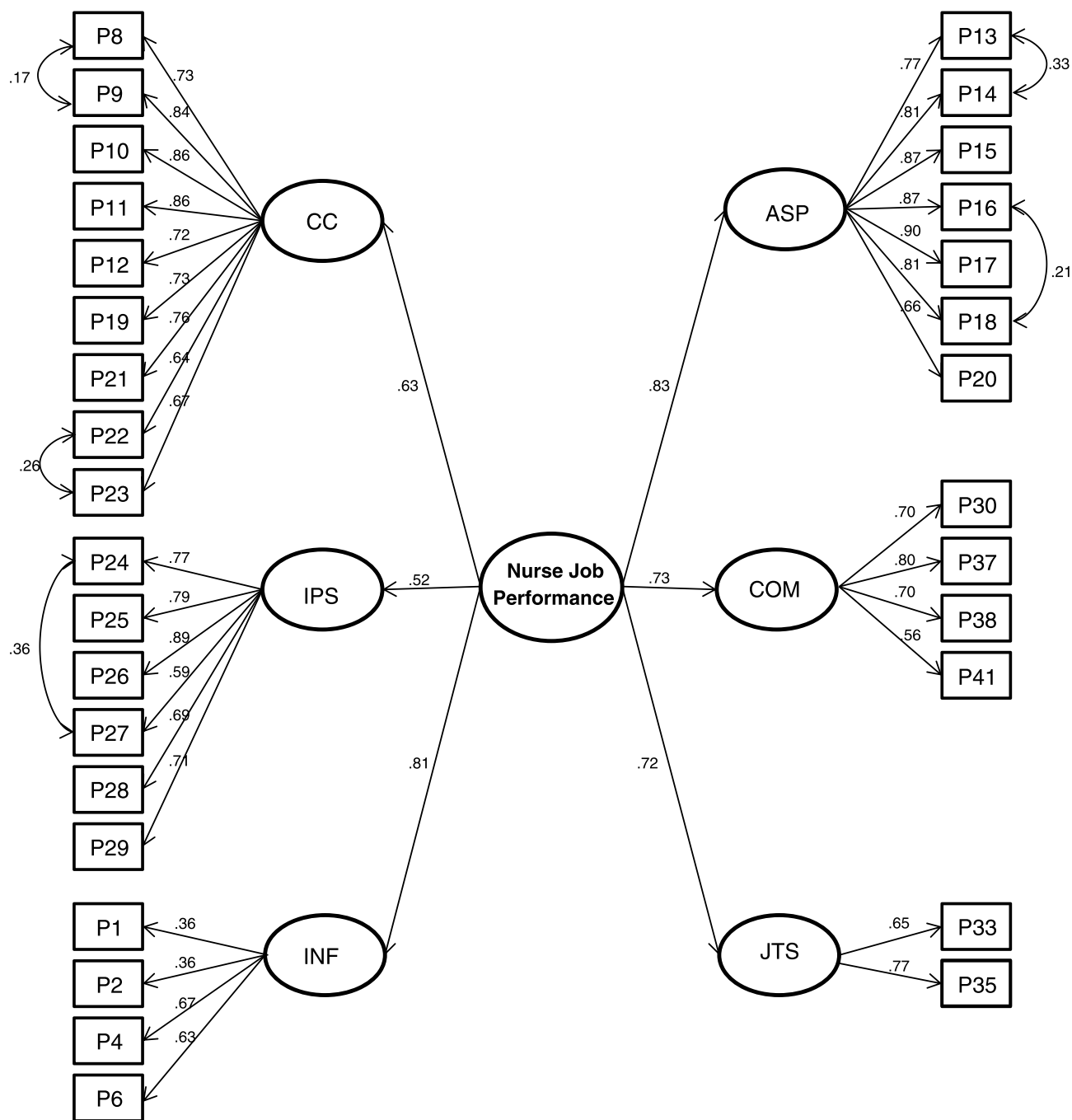


Figure 1. Confirmatory factor analysis results for the adapted version of the nurse Job Performance Scale. CC = coordination of care; ASP = assisting and supporting patients; IPS = interpersonal support; COM= compliance; INF = information; JTS = job task support.

of care” subscale in the TPS of the original scale were combined with Items 21–23 in the “technical care” subscale to create a new factor. This factor was called “coordination of care,” as in the original scale, because items containing the technical aspect of care were added. All five items in the “social support” subscale of the TPS and Item 20 in the “technical care” subscale of the original scale were included in the second factor. Different from the original scale, this subscale was called “assisting and supporting patients.” All items in the “interpersonal support” subscale in the CPS of the

original scale were included in the third factor, and the name of the factor was kept the same. One item from the “job task support” subscale of the CPS, two items from the “compliance” subscale, and one item from the “volunteering for additional duties” subscale were included in the fourth factor to create a new factor structure. This factor was called the “compliance” subscale by considering the compliance of the items it included. Items 1, 2, 4, and 6 in the TPS of the original scale were added to the fifth factor, and the name of the subscale was kept the same as “information,” as in

the original scale. Items 33 and 35 in the job task support subscale in the CPS of the original scale were added to the sixth factor. The name of the subscale was kept as “job task support,” as in the original scale.

It appears meaningful to emphasize the following points concerning the revised structure:

- The structure in TPS and CPS was preserved, although the items located in the subscales of both scales were moved to other subscales. For instance, the item “assisting patients with activities of daily living (e.g., showering, toileting, and feeding)” in the “technical care” subscale of the TPS in the original structure was also involved in the TPS in the Turkish version. However, the subscale in which it was involved changed and was moved to the “social support” subscale.
- Four of the five items in the subscale “technical care” of the original structure were involved in the subscale of “coordination of care,” as in the Turkish version. This condition may lead us to interpret that the nurses in Turkey evaluated planning, implementation, and technical activities in the context of nursing care as a whole.
- The item “assisting patients with activities of daily living (e.g., showering, toileting and feeding)” in the “technical care” subscale of the original structure was moved to the social support subscale in the Turkish version. Despite the fact that assisting patients to carry out daily life activities is a part of nursing care, it was considered that the shifting of this item to the latter scale was justified because the meaning of “assist” in Turkish is very close to the meaning of “support.”
- No change took place in the “interpersonal support” subscale that involved items pertaining to the communication and collaboration between the nurses in both the original and new structures. This may be due to the opinion that cooperation and support among nurses were realized in the Turkish sample because of their feeling of being professional colleagues.
- Three of the seven items on the “information” subscale in the original scale were omitted in the Turkish version, with four retained. The items on informing the patient and the family about discharge and out-of-hospital processes were preserved in the new version. However, it is remarkable that the items regarding care and treatment interventions administered during admission as well as the symptoms were not kept in the new version. In Turkey, physicians play the primary role in sharing information with the patients about admission procedures. Moreover, physicians tend to share information on symptoms, side effects, treatment, and care with the patients and their relatives together, rather than only with the patient (Atıcı, 2007).
- The number of items in the “compliance” subscale was three in the original scale and four in the Turkish version. Here, the item “complying with hospital rules, regulations and procedures, even when no one is watching” was omitted. In the original structure, the item “making

special arrangements for a patient’s family” involved in the job task support subscale of the original structure and the item “making innovative suggestions to improve the overall quality of the department” under the subscale of volunteering for additional duties were involved in the “compliance” subscale in the new version. Apparently, the matters of improving the related department and the making of special adjustments related to patient relatives by the sample in the Turkish structure were related to embracement to hospital, that is, the institution.

- In the new structure, only the items “making special arrangements for the patient” and “taking extra time to respond to a patient’s needs” remained under the subscale of “job task support.” The remaining three items were omitted. The item “making special arrangements for a patient’s family” in the subscale of job task support and the item “making innovative suggestions to improve the overall quality of the department” in the subscale of volunteering for additional duties in the original structure were involved in the subscale of “compliance” in the new version. This result is likely associated with reasons such as intense and long working hours and the care of excessive numbers of patients by a limited number of nurses. In addition to density of the current work of nurses, these items, including extra behaviors expected to be realized, were omitted.

Fourth stage: confirmatory factor analysis

Assigning an error covariance between items in line with modification suggestions is a method that has often been applied in the literature. However, higher values for error covariance signify that a model is less confirmative, although it does not invalidate the established model’s validity. What is important here is that the theoretical rationales for the assigned covariances are very explicitly ascribed (Şimşek, 2007). Five covariance assignments were conducted between the items that affected the model’s structure significantly and had theoretically similar meanings in this study (Figure 1). The first two items were “explaining to nurses in the unit the nature of the patient’s condition” and “reporting the critical elements of patients’ situations when turning over work shift” in the subscale “coordination of care.” As both items were related to patients’ clinical conditions, they were perceived similarly. The second two items were “showing care and concern to families” and “listening to families’ concerns” in the subscale “assisting and supporting patients.” The sample evaluated these two items as being close in meaning in both English and Turkish and thus interpreted them as related to care provided by the patient’s family. The third two items were “listening to patients’ concerns” and “showing care and concern to patients” in the subscale “assisting and supporting patients.” The concept of “concern” stood out in both items, with participants perceiving the items as having similar meanings. The fourth two

items were “administering medications and treatments” and “evaluating the effectiveness of nursing care” in the subscale “coordination of care.” Because nurses in Turkey do not consider the administration of medication and treatments as separate from nursing care, these two items were evaluated similarly by the sample. The last two items were “raising morale of other nurses in the unit” and “taking time to meet unit nurses’ emotional needs” in the subscale “interpersonal support.” These two items were related to emotional support between nurses. Therefore, the study sample perceived both items similarly.

After the corrections made during the CFA, the model showed a normal fit according to Şimşek (2007) in terms of CFI and acceptable fit in terms of χ^2/df , RMSEA, GFI, and AGFI (Table 2).

Fifth stage: internal consistency analysis

Cronbach’s alpha analyses are used in scale development studies to test the internal consistency of items using Likert scales (Polit & Beck, 2012). Coefficients that score between .80 and 1.00 indicate that the related scale has a high reliability (Tavşancıl, 2014). The Cronbach’s alpha reliability coefficient of the JPS in its internal consistency analysis was found to be .95 for the full scale (Table 3), showing that the scale items had a high internal consistency and internal reliability. In addition, the Cronbach’s alpha reliability coefficients were highly reliable for the first four subscales and were reliable for the fifth and sixth subscales (Table 3).

Sixth stage: criterion validity

This method, which is generally used in scale development studies and is also known as equivalent form reliability, is based on examining the results, obtained by applying a scale that is developed in two forms and with equivalent qualities to the same group incessantly at the same time, or intermittently at two different times, using Pearson correlation analysis (Gözüm & Aksayan, 2002). However, as developing equivalent test forms is difficult for cases in which a scale developed for the same purposes is present, the correlation between the present scale and the newly developed scale may be examined as an alternative approach (Gözüm & Aksayan, 2002).

Akgül (2005) emphasized that an absolute linear correlation must exist between the variables to calculate the Pearson correlation coefficient, suggesting that a scatter graph should be formed to examine the presence of a linear correlation between variables. In this study, the scatter graph of the measurements that were obtained from the JPS and EPS was examined, which revealed the presence of a positive linear correlation. A correlation that was positive, strong to moderate, and statistically significant was determined between the measurements obtained from the two scales in the analysis.

Conclusions and Recommendations

The scale that was developed by Greenslade and Jimmieson (2007) following the structure of Borman and Motowidlo

(1993), which approaches job performance as the two distinct domains of task performance and contextual performance, was not structurally valid in the Turkish context. The revised structure that was developed in this study, which used six subscales that related directly to performance with 32 items, was found to be a valid and reliable structure for a Turkish version of the scale.

The scope of the original English version of the scale seems to be a sufficient instrument for measuring the performance of nurses. However, the authors recommend testing scales in terms of validity and reliability before applying them in different cultures by considering that their structure is affected by intercultural differences.

Limitation

In conducting reliability analyses, a test–retest to show the scale’s reliability over time was not conducted.

Acknowledgment

We would like to thank the nurses who participated in the study.

Accepted for publication: May 15, 2016

*Address correspondence to: Rujnan Tuna, Faculty of Health Sciences, Department of Nursing, Unalan Mahallesi, Unalan Sokak, D-100 Karayolu Yanyol, 34700 Uskudar/Istanbul, Turkey.

Tel: +90 2162803153; E-mail: rujnantuna@yahoo.com

The authors declare no conflicts of interest.

Cite this article as:

Harmancı Seren, A. K., Tuna, R., & Eskin Bacaksız, F. (2018). Reliability and validity of the Turkish version of the Job Performance Scale instrument. *The Journal of Nursing Research*, 26(1), 27–35. doi:10.1097/jnr.0000000000000213

References

- Akgül, A. (2005). *Statistical analysis techniques in medical researches SPSS applications* (3rd ed.). Ankara, Turkey: Emek Ofset. (Original work published in Turkish)
- Atıcı, E. (2007). Factors affecting patient–physician relationship. *Uludağ Medical Journal*, 33(2), 91–96. (Original work published in Turkish)
- Bakker, A. B., Demerouti, E., & Euwema, M. C. (2005). Job resources buffer the impact of job demands on burnout. *Journal of Occupational Health Psychology*, 10(2), 170–180. doi:10.1037/1076-8998.10.2.170
- Behrenbeck, J. G., Timm, J. A., Griebenow, L. K., & Demmer, K. A. (2005). Nursing-sensitive outcome reliability testing in a tertiary care setting. *International Journal of Nursing Terminologies and Classifications*, 16(1), 14–20. doi:10.1111/j.1744-618X.2005.00002.x
- Borman, W. C., & Motowidlo, S. J. (1993). Expanding the criterion domain to include elements of contextual performance. In N. Schmitt & W. C. Borman (Eds.), *Personnel selection in organizations* (pp. 71–98). San Francisco, CA: Jossey-Bass.
- Borsa, J. C., Damásio, B. F., & Bandeira, D. R. (2012). Cross-cultural adaptation and validation of psychological instruments:

- Some considerations. *Paidéia*, 22(53), 423–432. doi:10.1590/S0103-863X2012000300014
- Büyüköztürk, Ş. (2011). *Manual of data analysis for social sciences* (15th ed.). Ankara, Turkey: Pegem Academy. (Original work published in Turkish)
- Coleman, V. I., & Borman, W. C. (2000). Investigating the underlying structure of the citizenship performance domain. *Human Resource Management Review*, 10(1), 25–44. doi:10.1016/S1053-4822(99)00037-6
- DeLucia, P. R., Ott, T. E., & Palmieri, P. A. (2009). Performance in nursing. *Reviews of Human Factors and Ergonomics*, 5(1), 1–40. doi:10.1518/155723409X448008
- Erdoğan, E. (2011). *The correlation between effective leadership, organizational silence, and performance* (Unpublished master's thesis). Gebze, Turkey: Gebze Advanced Technology Institute. (Original work published in Turkish)
- Gjersing, L., Caplehorn, J. R., & Clausen, T. (2010). Cross-cultural adaptation of research instruments: Language, setting, time and statistical considerations. *BMC Medical Research Methodology*, 10, 13. doi:10.1186/1471-2288-10-13
- Gözüm, S., & Aksayan, S. (2002). Guideline II for intercultural scale adaptation: Psychometric features and intercultural comparison. *HEMAR-G*, 4(2), 9–20. (Original work published in Turkish)
- Greenslade, J. H., & Jimmieson, N. L. (2007). Distinguishing between task and contextual performance for nurses: Development of a job performance scale. *Journal of Advanced Nursing*, 58(6), 602–611. doi:10.1111/j.1365-2648.2007.04256.x
- Gregg, A. C. (2002). Performance management data systems for nursing service organizations. *Journal of Nursing Administration*, 32(2), 71–78.
- Griffin, M. A., Neal, A., & Parker, S. K. (2007). A new model of work role performance: Positive behavior in uncertain and interdependent contexts. *Academy of Management Journal*, 50(2), 327–347. doi:10.5465/AMJ.2007.24634438
- Hamilton, K. E., Coates, V., Kelly, B., Boore, J. R., Cundell, J. H., Gracey, J., ... Sinclair, M. (2007). Performance assessment in health care providers: A critical review of evidence and current practice. *Journal of Nursing Management*, 15(8), 773–791. doi:10.1111/j.1365-2934.2007.00780.x
- Larrabee, J. H., Ostrow, C. L., Withrow, M. L., Janney, M. A., Hobbs, G. R. Jr., & Burant, C. (2004). Predictors of patient satisfaction with inpatient hospital nursing care. *Research in Nursing and Health*, 27(4), 254–268. doi:10.1002/nur.20021
- Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, 28(4), 563–575. doi:10.1111/j.1744-6570.1975.tb01393.x
- Öner, N. (2008). *Psychological tests used in Turkey: A reference guide* (2nd ed.). Istanbul, Turkey: Boğaziçi University Publishing House. (Original work published in Turkish)
- Pappas, S. H. (2008). The cost of nurse-sensitive adverse events. *Journal of Nursing Administration*, 38(5), 230–236. doi:10.1097/01.NNA.0000312770.19481.ce
- Pelletier, D., Duffield, C., Adams, A., Nagy, S., Crisp, J., & Mitten-Lewis, S. (2000). Australian nurse educators identify gaps in expert practice. *The Journal of Continuing Education in Nursing*, 31(5), 224–231. doi:10.3928/0022-0124-20000901-09
- Polit, D. F., & Beck, C. T. (2012). *Nursing research principles and method* (6th ed.). Philadelphia, PA: Lippincott Williams & Wilkins.
- Redfern, S. J., & Norman, I. J. (1990). Measuring the quality of nursing care: A consideration of different approaches. *Journal of Advanced Nursing*, 15(11), 1260–1271. doi:10.1111/j.1365-2648.1990.tb01741.x
- Schwirian, P. M. (1978). Evaluating the performance of nurses: A multidimensional approach. *Nursing Research*, 27(6), 347–351.
- Şencan, H. (2005). *Reliability and validity in social and behavioural measurements*. Ankara, Turkey: Seçkin. (Original work published in Turkish)
- Şimşek, Ö. F. (2007). *Introduction to the structural equation modeling: Principles and LISREL applications*. Ankara, Turkey: Ekinoks. (Original work published in Turkish)
- Tavşancıl, E. (2014). *Measurement of attitudes and data analysis with SPSS* (5th ed.). Ankara, Turkey: Nobel. (Original work published in Turkish)
- Wandelt, M., & Phaneuf, M. (1972). Three instruments for measuring the quality of nursing care. *Hospital Topics*, 50(8), 20–29. doi:10.1080/00185868.1972.9947925
- Wandelt, M. A., & Stewart Slater, D. (1975). *Slater nursing competencies rating scale*. New York, NY: Appleton-Century Croft.
- Whyte, D. A., Lugton, J., & Fawcett, T. N. (2000). Fit for purpose: The relevance of masters preparation for the professional practice of nursing: A 10-year follow-up study of postgraduate nursing courses in the University of Edinburgh. *Journal of Advanced Nursing*, 31(5), 1072–1080. doi:10.1046/j.1365-2648.2000.01388.x
- World Health Organization. (2008). *Process of translation and adaptation of instruments*. Retrieved from http://www.who.int/substance_abuse/research_tools/translation/en/