Measurement Invariance of the Chinese Job Performance Questionnaire (CJPQ) across Gender and Age among Doctors

CURRENT STATUS: UNDER REVIEW

Yanjie Yang

DOI: 10.21203/rs.3.rs-18133/v1

SUBJECT AREAS

Internal Medicine  Educational Philosophy and Theory

KEYWORDS

job performance, measurement equivalence, Chinese physicians
Abstract
Background At present, China urgently needs an effective and reliable tool to measure the job performance of doctors. For the first time, this study examined the measurement equivalence of the Chinese version of the Job Performance Questionnaire (CJPQ) in different genders and age groups among Chinese physician group.

Methods The subjects of this study were 1901 doctors from 7 ‘grade-three’ hospitals in Harbin, Heilongjiang Province, China, with an average age of 36.59±7.91 years. And this study adopted the Chinese version of the Job Performance Questionnaire (CJPQ) compiled by Chinese scholar Qiang Lin, which includes 3 dimensions, 33 items in total and adopts the five-point Likert scoring system. Demographic variables were analyzed by SPSS 21.0 and measurement consistency tests were performed using AMOS 17.0.

Results The results showed that for different genders, all changes in the CFI (Comparative Fit Index) were <0.01, indicating that the Chinese version of the Job Performance Questionnaire (CJPQ) is equivalent in terms of gender. Similarly, for subjects of different ages, all changes in CFI were also <0.01, indicating that the CJPQ is also equivalent in regard to age.

Conclusions The measurement equivalence of the Chinese version of the Job Performance Questionnaire (CJPQ) among Chinese physicians of different genders and age groups is fully established.

Background In organizational psychology, the most important dependent variable may be job performance. Job performance is defined as a terminology used to refers to scalable actions, behavior and outcomes that employees engage in or bring about that are linked with and contribute to organizational goals [1]. Many training programs are designed to improve job performance. Assessments of individuals are undertaken to identify their strengths and weaknesses in order to design training programs as well as for optimal placement decisions [2]. Performance appraisal, feedback and even merit pay systems make use of employee performance information. In short, job performance is a construct that is central to much of work psychology.
In China, with the rapid economic development since 1990s, the work pressure of employees in all walks of life has generally increased. At the same time, society has also put forward higher requirements for the service quality of the service industry. Especially in the field of medical services, Chinese doctors face greater stress. The ratio of doctors of China to general population is 1:735, which is considerably lower than that in development countries (1:280–1:640) [3]. The gap between healthcare demand and supply has led to an overworked state among Chinese doctors. Doctors, have become increasingly accountable for their actions and a clear need has arisen to recognize underperformance earlier and in a structured and transparent manner [4]. Consequently, there is a high demand for developing valid and reliable instruments to measure performance among doctors. The most widely used scale for measuring job performance is the Performance Questionnaire developed by Van Scotter & Motowidlo (1994). The scale has a total of 33 items, using Likert grade 5 scoring [5].

The Chinese Job Performance Questionnaire (CJPQ) was compiled by Qiang Lin et al according to the actual situation in China [6]. A total of 33 projects were identified, and it is characterized by task performance, interpersonal promotion and work dedication. Among them, there are 9 projects for task performance, 10 projects for interpersonal promotion, and 14 projects for work dedication. We use Likert 5-point scoring. 1 means “never” and 5 means “always”. The test reliability is good $\alpha=0.894$. A study by Huo Qi et al. (2014) has provided some evidences of the adequate psychometric properties for implementing the CJPQ. After verification, the Cronbach $\alpha$ coefficients of all dimensions were task performance 0.656, interpersonal promotion 0.706, and work dedication 0.858. The degree is 0.885 and the reliability is good. In 33 projects, the lowest value of correlation coefficient between each item and the corresponding dimension is 0.479 and the highest value is 0.729. The correlation coefficient between other items and their corresponding dimensions is distributed between 0.50-0.70. Structural validity is good. [7]

However, the problem is there was no research examining the CJPQ’ cross- sex and cross-age in doctors. Previous studies have shown that job performance is influenced by gender, age. Gao believed that sex was correlated with job performance [8]. This correlation might depend on the fact
that in the same organization, male and female shoulder different responsibilities and have different personal prospects, thus the variation in compensation and personal expectations. Age was another correlated with job performance scores. The level of job performance is in positive correlation with the age [8]. This is probably because employees can better adapt to the actual working environment and have lower expectations as he grows older. The above-mentioned conclusions have been verified in China, although there are some contradictory results [9, 10]. Samples from different working groups might account for the different results found across the studies; however, none of the questionnaires for measuring job performance in these studies has been tested for measurement invariance. Measurement invariance is a prerequisite for making meaningful comparisons between different groups [11]. In particular, measurement invariance determines the levels of similarity for the measurement methods and indicators in the groups being compared. It is critical for explaining similarities and differences across groups. Therefore, measurement invariance plays a critical role in explaining the similarities and differences among the characteristics of the groups being compared. Without testing measurement invariance in sex and age, researchers cannot determine the degree of influence that sex and age have on job satisfaction. As such, it cannot be ensured that the CJPQ will be useful in comparing different groups of Chinese doctors.

The first purpose of this study is to confirm the three-factor structure of the CJPQ short-form in a representative sample of Chinese doctors. The second purpose is to analysis the measurement invariance of the CJPQ short-form across sex and age. As no research has examined the measurement invariance of the job performance instruments in the CJPQ, this study will contribute to further understanding of the structure of job performance as well as designing successful interventions to improve doctor job performance.

Methods

Participants

Harbin Medical Sciences University is the leading medical institution in Heilongjiang Province northeast China. Using a Cluster sampling approach, it has seven grade A hospitals dispersed throughout heilongjiang province with a total of 2100 registered doctors. Self-reported anonymous
questionnaires were distributed to doctors, and valid questionnaires were 1901 (valid rate=90.52%). These doctors who completed the questionnaire served as our study participants.

**The Chinese Job Performance Questionnaire (CJPQ)**

The Chinese Job Performance Questionnaire (CJPQ) consists of 33 items revised by Qiang.et [6]. The revised questionnaire was made up of three subscales: Task performance (9 items), Interpersonal promotion (10 items) and Work dedication (14 items). Participants were asked about their feelings on a 5-point scale, ranging from 1 (never) to 5 (always). For the CJPQ, an internal consistency level of 0.894, as well as a three-factor structure, was found [6].

**Correlation analysis between sociodemographic variables**

Correlation analyses of the five sociodemographic variables (sex, age, marital status, education, professional qualifications) were performed using SPSS 21.0 for Windows. Statistical significance was defined as a two-tailed p-value of <0.05.

**Measurement Invariance**

The data was analyzed using AMOS 17.0 for confirmatory factor analysis to achieve structural validity test and measurement equivalence test. The specific inspection steps are as follows:

(1) Using a single set of confirmatory factor analysis method to test the degree of model fit of the three-factor structure of the job performance scale in the total sample and each demographic group, thereby establishing a well-fitting single-group baseline model. It is generally believed that when RMSEA ≤ 0.08 and CFI ≥ 0.90, the model reaches an ideal fit level;[12] (2) The multi group comparative test of confirmatory factor analysis was used to test the measurement equivalence of the job performance scale between the different groups of demographic variables. The test of equivalence consists of the following four levels from low to high: the morphological equivalence (Model 1), that is, to test whether the composition of latent variables in different groups is the same; the weak equivalence (Model 2), that is, whether the load of test factors are equal in each group; the strong equivalence (Model 3), that is, whether the intercepts of the observed variables are equal; the strictly equal (Model 4), that is, to test whether the error variance between different groups is equal. The four levels of equivalence tests have a hierarchical relationship and the tests must be performed
The measurement equivalence test is achieved by comparing the difference between the baseline model and the nested model. Therefore, in addition to the commonly used fitting indexes, such as $X^2$, TLI, CFI, and RMSEA, an incremental fitting indicator that reflects the difference between models is also needed. One commonly used indicator is $\Delta X^2$, which follows the distribution of degrees of freedom $\Delta df$; the other is $\Delta CFI$. Therefore, this study uses:

Two model fitting indicators: (1) RMSEA (approximate error root mean square); (2) CFI (relative fitting index). According to Hu and Bentler, Blackburn, and Donelly, it is required that each index meet the following criteria: RMSEA $< 0.08$, and the CFIs are all close to 0.90; the smaller the RMSEA is, the larger the CFI is, which indicates that the model fits better.

Two model comparison indicators: (1) $\Delta X^2$ value; (2) $\Delta CFI$ value. As with the $X^2$ test, the significant $\Delta X^2$ value indicates that there is a significant difference between the two models. The nested model should be rejected and the equivalent condition set by the nested model should not be accepted. The non-significant $\Delta X^2$ value indicates that the difference between the two models is not significant.

According to the principle of thrifty, a more concise nested model should be accepted and further restrictions can be imposed. Since the $\Delta X^2$ value is easily affected by the sample size, in the case of large samples, even if the model and the observation matrix fit well, the $\Delta X^2$ is easily significant, thus rejecting the correct theoretical model. According to Cheung and Rensvold, this study used $\Delta CFI$ as a supplemental model comparison indicator. $\Delta CFI < 0.01$ can be considered as no significant difference between the two models. In this study, the significance level was set as $P < 0.01$.

All data analysis is done by the AMOS 17.0 program operation.

Results

Demographic characteristics

Table 1 Participants’ characteristics
| Variables               | N(%)       |
|------------------------|------------|
| Sex                    |            |
| Male                   | 960(50.5%) |
| Female                 | 941(49.5%) |
| Age (years)            |            |
| ≤30                    | 507(26.7%) |
| 31-40                  | 911(47.9%) |
| >40                    | 483(25.4%) |
| Marital status         |            |
| Single                 | 563(29.6%) |
| Married/cohabiting     | 1338(70.4%)|
| Education              |            |
| College                | 410(41.6%) |
| Postgraduate or above  | 1491(78.4%)|
| Professional qualifications |        |
| None                   | 210(11%)   |
| Primary title          | 348(18.3%) |
| Intermediate title     | 606(31.9%) |
| Senior vice title      | 465(24.5%) |
| Senior title           | 272(14.3%) |

Table 1 shows the demographic characteristics of the subjects. A total of 1,901 doctors completed the measurement of the job performance scale. Among them, there are 960 (50.5%) men and 941 (49.5%) women. The average age of participants in this study was 36.59±7.91 years. Distribution by age: ≤ 30 years old (n=507), 31 to 40 years old (n=911), and >41 years old (n=483). 70.4% of participants are married or living together. 41.6% are university graduates, and 78.4% are masters or above.

**Correlation analysis**

**Table 2 Correlations between sex, age, marital status, education, professional qualifications**

| Variables               | 1   | 2  | 3  | 4  | 5   |
|------------------------|-----|----|----|----|-----|
| 1. Sex                 | 1.00|    |    |    |     |
| 2. Age (years)         | -.104** | 1.00|    |    |     |
| 3. Marital status      | -.084** | .468** | 1.00|    |     |
| 4. Education           | -.067** | .199** | .174** | 1.00|     |
| 5. Professional qualifications | -.089** | .707** | .493** | .223** | 1.00|

Table 2 shows the correlation between these five variables. Gender was negatively correlated with age, marital status, education level, and job title, with significant p-values. These results show that older doctors are more likely to be males, older female doctors are more likely to be single, and male doctors are more educated and job titles. Age is positively correlated with marital status, education level and job title. These results show that older doctors generally have more stable marital status, higher academic qualifications, and job titles than young doctors. Marital status is positively related to education level and professional title. The degree of education is positively related to the professional title.
Measurement Invariance

In this study, four demographic variables of job performance, including gender, age, education level, and professional qualifications, were measured for equivalence tests. Each factor measurement included five models: structural model, factor load invariant model, Project Intercept Invariant Model, Project Residual Invariant Model, and Structural Variance Invariant Model. Table 3 shows the results of the invariance test.

Table 3 Goodness-of-fit statistics for five invariance models

| Model                          | $\chi^2$  | $\Delta\chi^2$ | df | $\Delta$df | CFI     | $\Delta$CFI | RMSEA | RMSEA [90% CI] |
|-------------------------------|---------|----------------|----|-----------|--------|----------|------|----------------|
| **Sex**                       |         |                |    |           |        |          |      |                |
| Male (n = 960)                | 2598.529| —              | 492| —         | .900   | —        | .067 | [.064, .069]   |
| Female (n = 941)              | 2437.617| —              | 492| —         | .898   | —        | .065 | [.062, .067]   |
| Multiple group                |         |                |    |           |        |          |      |                |
| Configural model              | 5036.145| —              | 984| —         | .899   | —        | .047 | [.045, .048]   |
| Factor loadings invariant model| 5052.879| 15.091        | 1014| 33        | .900   | .000     | .046 | [.045, .047]   |
| Item intercepts invariant model| 5058.970| 6.091         | 1047| 33        | .900   | .000     | .045 | [.044, .046]   |
| Item residual variance invariant model| 5077.840| 18.870        | 1080| 33        | .901   | .001     | .044 | [.043, .045]   |
| Structural covariances invariant model| 5084.884| 7.044         | 1086| 6         | .901   | .000     | .044 | [.043, .045]   |
| **Age**                       |         |                |    |           |        |          |      |                |
| ≤30 (n = 507)                 | 1569.864| —              | 492| —         | .896   | —        | .066 | [.062, .069]   |
| 31–40 (n = 911)               | 2798.198| —              | 492| —         | .877   | —        | .072 | [.069, .074]   |
| >40 (n = 483)                 | 1736.360| —              | 492| —         | .890   | —        | .072 | [.069, .076]   |
| Multiple group                |         |                |    |           |        |          |      |                |
| Configural model              | 6104.553| —              | 1476| —        | .885   | —        | .041 | [.040, .042]   |
| Factor loadings invariant model| 6172.270| 67.717       | 1536| 66       | .885   | .002     | .040 | [.039, .041]   |
| Item intercepts invariant model| 6321.020| 148.750      | 1602| 66       | .883   | .000     | .039 | [.038, .040]   |
| Item residual variance invariant model| 6408.253| 87.233       | 1668| 66       | .883   | .001     | .039 | [.038, .040]   |
| Structural covariances invariant model| 6431.415| 23.162       | 1680| 12       | .882   | .001     | .039 | [.038, .040]   |

Note: $\chi^2$=Chi-square; df=degrees of freedom; CFI=confirmatory fit index; RMSEA=root mean square error of approximation; CI=90% Confidence Interval.

*indicates statistically-significant difference between nested models based on $\Delta$CFI>|.010|.

Sex

The initial model shows that the Chinese version of the job performance scale applies to the two groups of men and women. The structural covariance invariant model includes these two groups (CFI
The results of this study report the chi-square for all invariant models (5036.145, 5052.879, 5058.970, 5077.840, 5084.884), and also report the df for all invariant models (984, 1014, 1047, 1080, 1086). Chi-squared differences between the two nested models have also been reported (15.091, 6.091, 18.870, 7.044). All changes in CFI were <0.01, indicating that the Chinese version of the job performance scale is invariant across gender.

**Age**

The results show that in the three groups of ≤ 30 years old, 31-40 years old and > 40 years old, the CFI value of the initial index of the Chinese version of the job performance scale was less than 0.90. The structural covariance invariant model contains these three groups (CFI = 0.882; RMSEA = 0.039). The results of this study report the chi-squares for all invariant models (6104.553, 6172.270, 6321.020, 6408.253, 6431.415) and also report the df for all invariant models (1476, 1536, 1602, 1668, 1680). Chi-squared differences between two nested models were also reported (67.717, 148.750, 87.233, 23.162). All changes in CFI were <0.01, indicating that the Chinese version of the job performance scale is invariant across age.

**Tests for latent mean differences**

Finally, we compared latent mean differences in groups by fixing the latent mean values of the males group and the ≤30 group to zero and making a free estimate of the latent means of other groups. The comparison of the sex groups on the latent means revealed no statistically significant differences in the latent means of three factors of the CJPQ. The latent factor means were invariant in the sex groups, revealing that there were no statistically significant differences between males and females on all three subscale scores of the CJPQ. The same result also appeared in the age groups. These results indicated that, on average, there was no difference in the scores of all three subscales of the CJPQ.

**Discussion**

The highest level test is the structural covariances invariant model that the required factor variances and factor covariances are equal in the subgroups. All alternative values (i.e., CFI, RMSEA, and ΔCFI) of this study indicates good fit for the invariance criteria in sex and age [13] [14] [15] Because
measurement invariance is the prerequisite for making meaningful comparisons in different groups [16]. When all the above values support the criterions of measurement invariance, researchers can use the CJPQ to compare the job performance in different groups. It is important and meaningful to understand the differences between the sex groups and age groups. The measures can be adjusted for interventions, which are aimed at increasing the job performance levels of Chinese doctors.

There are statistically significant differences in gender and age in job performance of the CJPQ. For the variable sex, males tend to score higher than females in the job performance score. This is consistent with previous research results. [8] This might depend on the fact that in the same organization, male and female shoulder different responsibilities and have different personal prospects, thus the variation in compensation and personal expectations. However, some other studies find there is no difference between males and females for the score. [17, 18]

For the variable age, the job performance is positively correlated with age. There are similar conclusions in previous Chinese studies. [8] [9] [10] The opposite conclusion still exist, and several studies have found that younger doctors have the same score as the old ones. [18] There are some reasons for these different findings. One explanation is that employees can better adapt to the actual working environment and have lower expectations as he grows older. Another explanation is that the ratio of gender and age varies greatly in different hospitals. Therefore, we suggest that other researchers should be tested for invariance before conducting research into the sex and age effects on job performance.

**Restrictions and inadequacy**

Some possible limitations of this study should be interpreted. First of all, participants in this study were limited, to Heilongjiang Province; accurate representation of our study population may not be complete. Multicenter studies are needed to further confirm all the conclusions from the current study. Secondly, the study is conducted in large general hospitals. As there is different demand for the patient who went to different grades hospitals, doctors from different grades hospitals have huge differences in work categories and the proportion of staff. Therefore, the results in this study may not represent the job performance levels perceived by doctors from different grades hospitals and
community health centers. Thirdly, only self-report forms are adopted in this study which may have bias.

Future research will involve the multicenter studies to determine the measurement invariance of the CJPQ. We need to reply the evidence on the factor structure of the CJPQ for doctors come from different regions in future studies. At the same time, it is important to understand the nature of job performance in different grades hospitals and community health centers as well as incorporate the new psychological surveying with other self-reports. We hope that our study will further strengthen the awareness of this measurement invariance and gain insight into the validity and reliability of the CJPQ.

Conclusions
The results showed that for different genders, all changes in the CFI (Comparative Fit Index) were <0.01, indicating that the Chinese version of the Job Performance Questionnaire (CJPQ) is equivalent in terms of gender. Similarly, for subjects of different ages, all changes in CFI were also <0.01, indicating that the CJPQ is also equivalent in regard to age. The measurement equivalence of the Chinese version of the Job Performance Questionnaire (CJPQ) among Chinese physicians of different genders and age groups is fully established.

Declarations

Ethics approval and consent to participate
Local ethics committee ruled that no formal ethics approval was required in this particular case
Informed consent was obtained from all participants (veterinary officers and farmers)

Consent for publication
Not applicable

Availability of data and material
The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.
Competing interests

The authors declare that they have no competing interests

Funding

Not applicable

Authors' contributions

I, YY, am the only contributor in writing the manuscript. And I read and approved the final manuscript.

Acknowledgements

Not applicable

References

1. Chockalingam Viswesvaran DSO: Perspectives on Models of Job Performance. *International Journal of Selection and Assessment* 2000, Volume 8 Number 4
   December 2000:216-226.

2. Guion RM: Assessment, Measurement, and Prediction for Personnel Selection, 2 edn: Bowling Green State University; 1998.

3. Wang J-N SW, Chi T-S, Wu H, Wang L: Prevalence and associated factors of depressive symptoms among Chinese doctors a cross-sectional survey. *Int Arch Occup Environ Health* 2010, 83:905–911.

4. Southgate L CJ, David T: The assessment of poorly performing doctors the development of the assessment programme for the GMC’s performance procedure. *Med Educ* 2001, 35:2–8.

5. Scotter SJMaJRV: Evidence That Task Performance Should Be Distinguished From Contextual Performance. *Journal of Applied Psychology* 1994, Vol. 79.No. 4.:475-480.
6. QiangLin: Research on the relationship between organizational commitment, psychological capital and job performance. Tianjin Normal University; 2009.

7. Qi H: Research on the relationship between work happiness, trait anxiety and job performance of health care workers. Henan University; 2014.

8. Shuyan G: Research on the Impact of Organizational Commitment and Organizational Support on Employees' Work Performance. Shanghai Jiaotong University; 2007.

9. Min L: Learning style of primary and secondary school teachers and its relationship with role pressure and job performance. Hangzhou Normal University; 2011.

10. Lijun X: Study on the Relationship among Emotional Labor, Emotional Intelligence and Job Performance of Preschool Teachers. Hunan Normal University; 2012.

11. Vandenberg R J LCE: A Review and Synthesis of the Measurement Invariance Literature: Suggestions, Practices, and Recommendations for Organizational Research. *Organizational Research Methods* 2000, 2000, 5(3):4-69.

12. Hu Lt, Bentler PM: Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal* 1999, 6(1):1-55.

13. Kline RB: Principles and Practice of Structural Equation Modeling. 2011, 1:427.

14. Cheung GW RR: Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance. *Structural Equation Modeling A Multidisciplinary Journal* 2002, 9(2):233-255.

15. Browne MW CR, Bollen KA, Long JS: Alternative ways of assessing model fit. *Sage focus editions* 1993, 154:136-136.

16. Schmidt P BT: Rezension Timothy A. Brown, 2006 Confirmatory factor analysis for applied research; 2010.

17. Xuefen C: Relationship between personality, coping style and job performance of
middle school teachers. Fujian Normal University; 2007.

18. Xiaoli M: The relationship between employees' emotional intelligence, achievement motivation and job performance. East China Normal University; 2011.