ORIGINAL RESEARCH

# The Chinese-Adapted Jefferson Scale of Empathy -Medical Student Version for Assessing the Level of Empathy in Medical Students: A Study of the Reliability and Validity

liong Liu<sup>1,\*</sup>, Tonghao Lin<sup>1,\*</sup>, Wencheng Wu<sup>1</sup>, Hong Duo<sup>2</sup>, Cuijiao Tian<sup>3</sup>

Department of Radiation and Medical Oncology; Hubei Key Laboratory of Tumor Biological Behaviors; Hubei Province Cancer Clinical Study Center, Zhongnan Hospital, Wuhan University, Wuhan, Hubei, People's Republic of China; <sup>2</sup>Zhongnan Hospital of Wuhan University, Institute of Hepatobiliary Diseases of Wuhan University, Transplant Center of Wuhan University, Hubei Key Laboratory of Medical Technology on Transplantation, Wuhan, 430071, People's Republic of China; <sup>3</sup>Teaching Affairs Office, Zhongnan Hospital of Wuhan University, Wuhan, Hubei, 430071, People's Republic of China

\*These authors contributed equally to this work

Correspondence: Cuijiao Tian, Email ZN000985@whu.edu.cn

Purpose: To investigate the psychometric properties of the Chinese-Adapted Jefferson Scale of Empathy-Medical Student Version (JSE S-Version) and evaluate its applicability.

Methods: The JSE S-Version was translated into Chinese and subsequently converted into an electronic questionnaire. Then 320 medical students were included in the survey by convenience sampling. The reliability of the scale was assessed by internal consistency reliability, and the validity of the scale was assessed by principal component analysis, exploratory factor analysis and validation factor analysis.

Results: Principal component analysis extracted three principal components. Exploratory factor analysis showed that the structure of the remaining 19 items was consistent with that of the original English version, except for item 18. Confirmatory factor analysis showed that the scale demonstrated better convergent validity and discriminant validity after deleting item 18. All items demonstrated both adequate item-total score correlations and favorable discrimination index effect sizes. The internal consistency reliability coefficients of the three dimensions of the Chinese version of the JSE-S scale were all greater than 0.8.

Conclusion: The Chinese-adapted JSE S-Version demonstrated robust psychometric properties consistent with the original English version, with satisfactory reliability and validity analyses, establishing its appropriateness for use in Chinese populations.

Keywords: the Jefferson scale of empathy, medical student version (JSE S-Version), reliability and validity, medical students, Empathy

#### Introduction

In 1948, the World Health Organization (WHO) first proposed the three-dimensional concept of health: "Health is a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity".<sup>1</sup> This model is consistent with The Bio-Psycho-Social model of disease in medicine. This model requires medical workers to not only focus on and treat the patient's physical illness, but also to pay attention to the patient's emotional feedback and psychological needs. By understanding and respecting patients' feelings, treatment activities can be carried out to promote the patient's well-being in the four dimensions of physiological, psychological, behavioral, and social functioning. Therefore, empathy has become an important quality for health care providers.

Hojat et al introduced the concept of empathy in the context of patient care, suggesting that empathy is primarily a cognitive (rather than affective or emotional) attribute that involves understanding patients' experiences, concerns, and perspectives (rather than feelings), the ability to convey this understanding, and a willingness to help.<sup>2</sup> Research shows

that empathy helps improve patient satisfaction, thereby promoting their compliance, reducing patient anxiety and depression, and is associated with improved clinical outcomes.<sup>3</sup> In addition, empathic physicians are more likely to achieve job satisfaction, psychological well-being, and make better clinical decisions.<sup>4–6</sup>

To date, the measurement of empathy remains an inescapable challenge. Various methods have been employed to assess medical students' empathy, including patient assessments, faculty or standardized patient assessments, self-assessments, and even neurobiological assessments.<sup>7–12</sup> Hojat et al developed a self-report instrument called the Jefferson Scale of Empathy (JSE), which uses a 7-point Likert scale and consists of 20 items across three dimensions. There are three versions of the scale: the Jefferson Scale of Empathy, Health Professions Version (JSE HP-Version) for physicians and other health professionals, the Jefferson Scale of Empathy, Medical Student Version (JSE S-Version) for medical students, and the Jefferson Scale of Empathy, Health Professions Student Version (JSE HPS-Version) for students in any health profession discipline other than medicine.<sup>2</sup> JSE S-Version is considered to be the most suitable tool for assessing the empathy skills of clinical medical students. This tool has a three-factor structure, including perspective taking, compassionate care, and the ability to stand in the patient's shoes.<sup>2,13,14</sup>

The JSE S-Version has received exceptionally high evaluations in cross-sectional studies on empathy ability conducted across multiple countries, and remains the most widely utilized empathy assessment scale worldwide. Since empathy evaluation can be influenced by comprehensive factors including social environment, culture, and education level, it becomes necessary to employ linguistically translated versions of the JSE S-Version in diverse sociocultural contexts to examine their psychometric properties. Previous research has demonstrated the scale's robust cross-cultural assessment capability, with most translated versions exhibiting similar latent variable structures.<sup>15–22</sup> Although a minority of researchers have attempted to adopt a four-factor structure,<sup>23,24</sup> the additional factors extracted through statistical methods typically demonstrate minimal variance loadings, which may compromise the reliability of the JSE-S-Version. Consequently, these alternative factor structures are generally disregarded in research practice.

Currently, there are more reports from scholars in China using JSE S-Version to study the empathy levels of medical students and its influencing factors.<sup>25,26</sup> These studies primarily focus on measuring empathy within specific populations and identifying objective determinants affecting empathy in medical students. Research findings consistently demonstrate significant associations between Chinese medical students' empathy levels and factors such as age, educational attainment, and family environment. A particularly concerning finding reveals a paradoxical decline in overall empathy levels among medical students during their basic medical education phase, with recovery only observed upon transitioning to clinical training.<sup>27,28</sup> This pattern suggests potential deficiencies in role-immersion components within China's preclinical medical curriculum. To establish a comprehensive empathy education system, it is imperative to develop standardized empathy assessment tools.

Despite the growing utilization of the JSE S-Version in China, there remains a notable paucity of psychometric validation studies for its Chinese version. Existing reports on the reliability and validity of Chinese JSE adaptations are outdated, and no English-language publications examining the psychometric properties of the JSE S-Version can be retrieved from PubMed. Consequently, the validity and applicability of the JSE S-Version in Chinese populations remain questionable, resulting in insufficient evidence to support the use of its Chinese version in clinical and research settings in China. Given the absence of English-language psychometric validation studies for the Chinese-adapted JSE-S-Version, this study aims to: (a) determine whether the latent factor structure of the Chinese-adapted JSE S-Version aligns with that of the original English version. (b) systematically evaluate its reliability and validity metrics, and (c) assess its suitability as an empathy assessment tool for Chinese medical students. To address these objectives, we will conduct a comprehensive survey study utilizing the Chinese JSE S-Version among medical students at Wuhan University.

## **Materials and Methods**

#### Subjects

The participants in the study are 320 undergraduate medical students from Wuhan University, including 84 third-year students (Class of 2020), 105 fourth-year students (Class of 2019), and 131 fifth-year students (Class of 2018).

#### Survey Instrument

This study used the Chinese-adapted JSE-S-Version (Thomas Jeffferson University, 2001. All rights reserved) to assess the clinical empathy of undergraduate medical students. Before using the JSE-S-Version, we obtained permission from the Jefferson Medical College. The scale consists of three factors: "perspective taking", "compassionate care", and "standing in the patient's shoes", with a 7-point Likert scale rating system. It comprises 20 items, including ten positive and ten negative items. For positive items, responses are directly scored (strongly disagree=1, strongly agree=7). For negative items, responses are reverse-scored (strongly disagree=7, strongly agree=1). The total score is calculated, with a maximum score of 140, where higher scores indicate higher levels of empathy.

The researcher referred to the method of Hojat M and Erdmann JB,<sup>15</sup> two researchers with medical background and proficient English skills independently translated the original JSE S-Version into Chinese. The two translations were then compared and integrated to form the Chinese version. Two medical professors were invited to further discuss and revise the scale to get the initial Chinese version of the JSE-S. Subsequently, three medical English experts were invited to follow the back-transformation and coordination procedure to independently translate the first version of the Chinese JSE S-Version into the English version and compare it with the original scale. A pre-survey was conducted within a small sample, and the translation was fine-tuned according to the survey results to ensure that the content of the items, the main meanings, and the comprehensibility remained unchanged. This finalized version was then determined as the final Chinese-adapted JSE-S-Version.

## Sample Size Calculation

From the perspective of reliability and validity, the sample size should be at least 5–10 times of the number of items in the scale. Therefore, the sample size included in this study should be 100–200 people or more.

## **Research Methods**

Considering the feasibility of the experiment, we adopted a cross-sectional method and utilized a web-based questionnaire to conduct the survey. The questionnaire was divided into two parts, the first part was basic information such as gender, grade, major, and urbanization level of home address, and the second part was the Chinese-adapted JSE-S-Version. The online questionnaire was distributed to all students in the target grades on September 1, 2023. To ensure data integrity, each student account was restricted to a single submission, and all collected responses were anonymized by excluding personally identifiable information. Survey participation was automatically closed upon meeting either of the following criteria: (1) achieving a response rate of 60% of the total distributed questionnaires, or (2) reaching the 72-hour time limit (whichever occurred first). Statistical analysis was initiated immediately after data collection concluded. At this time, the Class of 2018 (fifth-year) students had graduated from their undergraduate program approximately two months earlier; the Class of 2019 (fourth-year) students were transitioning to their fifth year and had just begun their clinical internships; and the Class of 2018 (third-year) students were about to enter their fourth year.

The study was approved by the Ethics Committee of Zhongnan Hospital of Wuhan University prior to implementation. A senior administrator was appointed to oversee the survey administration. Prior to the start of the survey, students were informed of the study via an online campus bulletin, and anonymity was emphasized at all stages of the study. No personally identifiable information was collected and strict confidentiality measures were implemented for all personal data.

## Statistical Analysis

Mean  $\pm$  Standard Deviation was used to describe the quantitative data, and rates and composition ratios were used to describe the qualitative data. Internal consistency reliabilities (represented by Cronbach's  $\alpha$  coefficient and McDonald's  $\omega$  coefficient) were used to assess the reliability of the Chinese-adapted JSE-S-Version for entire scale and sub-structures. To examine the structural validity of the scale, principal component analysis (PCA) was first used, followed by random division of the sample into two equal subgroups - one for exploratory factor analysis (EFA) and the other for confirmatory factor analysis (CFA). Pearson correlation coefficients were used to examine the relationship between each item's score and the total JSE score. To evaluate the discriminative capacity of each item, the top one-third of

participants with the highest JSE scores and the bottom one-third with the lowest JSE scores were extracted from the whole sample. For each individual item, the discrimination index effect size was calculated by computing the mean score difference between the top -third and bottom -third JSE scoring groups and dividing this value by the item's pooled standard deviation.<sup>29</sup> The data were analyzed using SPSS25.0 and jamovi with a significance level set at  $\alpha$ =0.05.

## **Results**

## General Characteristics of the Study Subjects

A total of 320 questionnaires were collected in this study, with 84 (26.25%) from the Class of 2020, 105 (32.81%) from the Class of 2019 and 131 (40.94%) from the Class of 2018. There were 135 (42.19%) males and 185 (57.81%) females. Those who participated in the internship accounted for 244 (76.25%) and those who did not participate in the internship 76 (23.75%).248 (77.50%) had the intention to engage in medicine and 72 (22.50%) had no intention to engage in medicine (Table 1).

## Principal Component Analysis of the Chinese-Adapted ISE-S-Version

The Chinese-adapted JSE-S-Version is a hierarchical data with a Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy of 0.890 and Bartlett's test of sphericity resulting in  $\chi = 3249$  (P<0.001), indicating suitability for principal component analysis. Using the principal component analysis method without rotation, three principal components can be extracted based on eigenvalues greater than 1. The variance explained by the three principal components are 34.13%, 17.85%, and 5.88% respectively, with a cumulative explained variance of 57.9%. The factor loading coefficients of the 20 items of the empathy scale on each principal component are all >0.4. Except for the unique residual of 0.687 for item 15 in the empathy scale, the unique residuals of the remaining 19 items were all less than 0.6. This suggests that, except for entry 15, all the other items were extracted to a relatively good extent. (Supplementary Tables 1 and 2)

## Exploratory Factor Analysis for the Chinese-Adapted JSE-S-Version

The number of fixed factors was 3 and factor loading less than 0.3 were not shown. The method of minimum residual extraction combined with oblique rotation (oblimin) was used to extract and rotate the factors. The results show that Factor 1 contains item 2, 4, 5, 9, 10, 13, 15, 16, 17, 20, a total of 10 items; factor 2 contains item 1, 7, 8, 11, 12, 14, 19, a total of 7 items; factor 3 contains item 3, 6, a total of 2 items. Except for item 18, the factor loading coefficients of the remaining 19 items are greater than 0.5 in one factor, and less than 0.3 in the other two factors, no "confusion" phenomenon. Except for item 18, whose factor loadings in all three factors are less than 0.3, the classification results of the other 19 items are consistent with the original English version, with no "misclassification" phenomenon. The explanatory rates of the three principal component factors are all greater than 5%, and the cumulative variance explained

| Table I Population Statistical  | Characteristics     |
|---------------------------------|---------------------|
| Factor                          | Population (ratio%) |
| Class                           |                     |
| 2020                            | 84 (26.25)          |
| 2019                            | 105 (32.81)         |
| 2018                            | 131 (40.94)         |
| Gender                          |                     |
| Male                            | 135 (42.19)         |
| Female                          | 185 (57.81)         |
| Participate in the internship   |                     |
| Yes                             | 244 (76.25)         |
| No                              | 76 (23.75)          |
| Intention to engage in medicine |                     |
| Yes                             | 248 (77.50)         |
| No                              | 72 (22.50)          |
|                                 |                     |

is 51.95%. Due to the low factor loading of item 18 and its unique residual value of 0.867, after removing item 18, the factor dimension of the remaining items did not change, and the factor loadings were all greater than 0.5, with a cumulative variance explained of 53.93%. Considering that the unique residual values of item 15, 19, 5 are greater than 0.6, following the order of unique residual values from large to small, item 15, 19, 5 are successively removed. After removal, the factor loadings of all items are greater than 0.595, unique residual values are all less than 0.6, and the cumulative variance explained can reach 58.86%. The scree plot demonstrated that the eigenvalue curve flattened after extracting the third factor, supporting the three-factor structure of the Chinese-adapted JSE-S-Version (Figure 1).

#### Confirmatory Factor Analysis for the Chinese-Adapted JSE-S-Version

According to the item framework of the JSE-S-Version, confirmatory factor analysis was conducted. In terms of convergent validity analysis, all items showed significance at the 0.001 level. Only item 18 had standard loading coefficients of 0.283, which were less than 0.5, while the remaining 19 items were all greater than 0.5. After deleting the item 18, the significance of all items is less than 0.001, and the standard loadings are all greater than 0.5, indicating a good correspondence between factors and items overall (Figure 2). Modification indices of the remaining 19 items all meeting the standard of modification indices less than 15 (Supplementary Table 3). The Average Variance Extracted (AVE) values for the three factors are 0.493, 0.546, and 0.635 respectively, approximately meeting the requirement of AVE greater than 0.5. The Composite Reliability (CR) values for the three factors are 0.905, 0.891, and 0.777 respectively, all meeting the requirement of CR greater than 0.7. Although the Average Variance Extracted (AVE) of the first factor was slightly below 0.5, its composite reliability (CR) exceeded 0.7, suggesting acceptable convergent validity for this factor. Comprehensive evaluation of factor-item correspondence, factor loadings, modification indices, AVE values and CR coefficients confirms that the scale maintains satisfactory convergent validity. In the analysis of discriminant validity, the results show that the square root of the Average Variance Extracted (AVE) for the three factors are 0.702, 0.739, and 0.797, respectively, all of which are greater than the Pearson correlation coefficients between each factor and the other two factors, indicating good discriminant validity of the scale data (Supplementary Table 4). Furthermore, the Harman's single-factor test was conducted to examine common method bias, and no factor explains more than 50% of the variance, suggesting the absence of common method bias. Additionally, when all items were loaded onto a single factor for confirmatory factor analysis, the fit indices were all relatively poor, further confirming the absence of common method bias. (Supplementary Tables 5 and 6)

## Item-Total Score Correlation And Discrimination Index Effect Size

The item-total score correlations of each item ranged from 0.409 to 0.810. All correlations were positive and statistically significant (p < 0.001), indicating that each item contributes positively and significantly to the overall score of the JSE scale. The item discrimination effect size indices ranged from 0.921 to 1.893. Cohen suggests that the effect size values around 0.30 or lower are considered negligible, around 0.50 are moderate, and around 0.70 and higher are large and practically important.<sup>29</sup> The discrimination index effect sizes confirm that all items significantly discriminate students' empathy capabilities.

## Reliability Analysis of JSE S-Version

According to the factor structure of the original English version of the JSE S-Version, a reliability analysis was conducted on the Chinese translated version. The results showed that the Cronbach's  $\alpha$  coefficient of the total scale was 0.885, and McDonald's  $\omega$  coefficient was 0.896. After removing item 15, the overall Cronbach's  $\alpha$  of the scale increased to 0.888, and the McDonald's  $\omega$  coefficient increased to 0.897. Removing item 18 did not significantly change the overall internal consistency of the scale. However, removing the remaining 18 items would decrease the overall internal consistency coefficient of the scale. The Cronbach's  $\alpha$  coefficients of the three factors are 0.881, 0.877, and 0.821, while the McDonald's  $\omega$  coefficients are 0.894, 0.886, and 0.822, indicating good internal consistency of the three dimensions of the Chinese translated version of the JSE-S scale. Specifically, after removing item 15, the Cronbach's  $\alpha$  coefficient and McDonald's  $\omega$  coefficient of the first factor increased to 0.888 and 0.896; after removing item 18, the Cronbach's  $\alpha$ 

| item -     |               |                   |              |            |               | Factor     |               |              |            | E.               | actor                 |          |
|------------|---------------|-------------------|--------------|------------|---------------|------------|---------------|--------------|------------|------------------|-----------------------|----------|
|            | 1             | 2 3               | Uniqueness   | Item       | 1             | 2          | 3             | Uniqueness   | Item       | 1                | 2 3                   | Uniqu    |
| 20         | 0.843         |                   | 0.274        | 20         | 0.843         |            |               | 0.273        |            |                  |                       |          |
| 10         | 0.772         |                   | 0.453        | 10         | 0.772         |            |               | 0.451        | 20         | 0.839            |                       | 0        |
| 13         | 0.738         |                   | 0.447        | 13         | 0.738         |            |               | 0.444        | 10         | 0.774            |                       | 0        |
| 17         | 0.736         |                   | 0.385        | 17         | 0.737         |            |               | 0.385        | 13         | 0.766            |                       | 0        |
| 16         | 0.721         |                   | 0.385        | 16         | 0.722         |            |               | 0.384        | 17         | 0.730            |                       | 0        |
| 9          | 0.678         |                   | 0.559        | 9          | 0.678         |            |               | 0.559        | 16         | 0.714            |                       | 0        |
| 4          | 0.677         |                   | 0.486        | 4          | 0.677         |            |               | 0.486        | 9          | 0.690            |                       | 0        |
| 2          | 0.655         |                   | 0.568        | 2          | 0.656         |            |               | 0.568        |            | 0.090            |                       | 0        |
| 5          | 0.568         |                   | 0.707        | 5          | 0.569         |            |               | 0.705        | 4          | 0.690            |                       | 0        |
| 15         | 0.542         |                   | 0.737        | 15         | 0.542         |            |               | 0.738        | 2          | 0.635            |                       | 0        |
| 7          | C             | 0.852             | 0.146        | 7          |               | 0.843      |               | 0.139        | 7          | 0                | .815                  | 0        |
| 12         | C             | 0.791             | 0.479        | 12         |               | 0.798      |               | 0.464        | 12         | 0                | .809                  | 0        |
| 11         | C             | 0.769             | 0.386        | 11         |               | 0.763      |               | 0.384        | 11         | 0                | 768                   | 0        |
| 8          |               | 0.709             | 0.328 -      | 8          |               | 0.705      |               | 0.318        |            | 0                | 698                   | 0        |
| 14         |               | 585               | 0.421        | 14         |               | 0.679      |               | 0.424        | 0          | 0                | .090                  |          |
| 19         | C             | 1.569             | 0.585        | 1          |               | 0.591      |               | 0.590        | 14         | 0                | .671                  | 0        |
| 3          |               | 0 791             | 0.307        | 19         |               | 0.557      |               | 0.687        | 1          | 0                | .599                  | C        |
| 6          |               | 0.730             | 0.399        | 3          |               |            | 0.738         | 0.380        | 6          |                  | 0.739                 | 0        |
| 18         |               |                   | 0.867        | 6          |               |            | 0.724         | 0.374        | 3          |                  | 0.737                 | 0        |
| Table1-2 E | Explained var | iance ratio(items | not deleted) | Table2-2 I | Explained var | iance rati | o(delete iter | n 18)        | Table3-2 I | Explained variar | nce ratio(delete item | 18,15,19 |
| Factor     | SS Loadings   | % of Variance     | Cumulative % | Factor     | SS Loading    | js %o      | f Variance    | Cumulative % | Factor     | SS Loadings      | % of Variance         | Cumu     |
| 1          | 4.987         | 24.935            | 24.935       | 1          | 4.988         |            | 26.253        | 26.253       | 1          | 4.426            | 27.663                | ā        |
| 2          | 3.869         | 19.346            | 44.281       | 2          | 3.838         |            | 20.200        | 46.453       | 2          | 3.502            | 21.888                | 4        |
| 3          | 1 534         | 7 670             | 51 951       | 3          | 1 420         |            | 7 475         | 53 928       | 3          | 1.442            | 9.010                 | i.       |
| Scree plot | 1             |                   |              | Scree plot | 2             |            |               |              | Scree plot | 3                |                       |          |

Figure I Exploratory factor analysis process diagram.

Liu et al

| CFA (items not deleted) |       |          |       |        |        |          |             |                | CFA (delete item 18) |           |          |       |        |        |                 |  |
|-------------------------|-------|----------|-------|--------|--------|----------|-------------|----------------|----------------------|-----------|----------|-------|--------|--------|-----------------|--|
| Fastar                  | ltaus | Fatimate | cr.   | 7      | _      | Stand.   | ltem-total  | Discrimination | Factor               | Indicator | Estimate | SE    | z      | р      | Stand. Estimate |  |
| Factor 1                | 2     | Estimate | SE    | Z      | p      | Estimate | correlation | Index effect   | Factor 1             | 2         | 0.766    | 0.086 | 8.934  | < .001 | 0.650           |  |
|                         |       | 0 7 7 7  | 0.000 | 0.020  |        |          |             | 1 204          |                      | 4         | 0.624    | 0.063 | 9.876  | < .001 | 0.702           |  |
| Factor I                | 2     | 0.767    | 0.060 | 0.930  | < .001 | 0.050    | 0.542       | 1.394          |                      | 5         | 0.627    | 0.092 | 6.846  | < .001 | 0.523           |  |
|                         | 4     | 0.624    | 0.063 | 9.874  | < .001 | 0.702    | 0.557       | 1.379          |                      | 9         | 0 732    | 0 079 | 9 239  | < 001  | 0.667           |  |
|                         | 5     | 0.627    | 0.092 | 6.847  | < .001 | 0.523    | 0.553       | 1.518          |                      | 10        | 0.732    | 0.075 | 10.102 | 0.001  | 0.310           |  |
|                         | 9     | 0.732    | 0.079 | 9.239  | < .001 | 0.667    | 0.602       | 1.233          |                      | 10        | 0.743    | 0.073 | 10.192 | < .001 | 0.718           |  |
|                         | 10    | 0.743    | 0.073 | 10.191 | < .001 | 0.718    | 0.417       | 1.127          |                      | 13        | 0.660    | 0.062 | 10.702 | < .001 | 0.743           |  |
|                         | 13    | 0.660    | 0.062 | 10.701 | < .001 | 0.743    | 0.560       | 1.263          |                      | 15        | 0.680    | 0.099 | 6.856  | < .001 | 0.523           |  |
|                         | 15    | 0.681    | 0.099 | 6.858  | < .001 | 0.523    | 0.810       | 1.893          |                      | 16        | 0.729    | 0.063 | 11.490 | < .001 | 0.782           |  |
|                         | 16    | 0.729    | 0.063 | 11.491 | < .001 | 0.782    | 0.737       | 1.767          |                      | 17        | 0.646    | 0.056 | 11.634 | < .001 | 0.788           |  |
|                         | 17    | 0.646    | 0.056 | 11.635 | < .001 | 0.788    | 0.519       | 1.445          | <b>→</b>             | 20        | 0.750    | 0.058 | 13 019 | < 001  | 0.848           |  |
|                         | 20    | 0.750    | 0.058 | 13.019 | < .001 | 0.848    | 0.551       | 1.197          | ,                    | 20        | 0.750    | 0.050 | 15.015 | < .001 | 0.040           |  |
| Factor 2                | 1     | 1.019    | 0.120 | 8.523  | < .001 | 0.622    | 0.708       | 1.751          | Factor 2             | 1         | 1.017    | 0.120 | 8.504  | < .001 | 0.620           |  |
|                         | 7     | 1.167    | 0.075 | 15.595 | < .001 | 0.941    | 0.606       | 1.631          |                      | 7         | 1.171    | 0.075 | 15.712 | < .001 | 0.945           |  |
|                         | 8     | 1.145    | 0.090 | 12.724 | < .001 | 0.831    | 0.596       | 1.553          |                      | 8         | 1.145    | 0.090 | 12.724 | < .001 | 0.831           |  |
|                         | 11    | 1.146    | 0.101 | 11.370 | < .001 | 0.771    | 0.714       | 1.693          |                      | 11        | 1.145    | 0.101 | 11.371 | < .001 | 0.771           |  |
|                         | 12    | 1.081    | 0.127 | 8.526  | < .001 | 0.625    | 0.409       | 1.249          |                      | 12        | 1.072    | 0.127 | 8.440  | < .001 | 0.620           |  |
|                         | 14    | 1.092    | 0.101 | 10.797 | < .001 | 0.745    | 0.673       | 1.745          |                      | 14        | 1 087    | 0 101 | 10 734 | < 001  | 0 742           |  |
|                         | 18    | 0.438    | 0 124 | 3 521  | < 001  | 0.283    | 0.668       | 1 825          |                      | 10        | 1.007    | 0.101 | 7.024  | 0.001  | 0.742           |  |
|                         | 19    | 0.881    | 0.116 | 7 607  | < 001  | 0.205    | 0.425       | 1 141          |                      | 19        | 0.883    | 0.116 | 7.634  | < .001 | 0.568           |  |
| Eactor <sup>2</sup>     | 2     | 1.246    | 0.110 | 0.755  | < 001  | 0.500    | 0.507       | 0.021          | Factor 3             | 3         | 1.244    | 0.128 | 9.703  | < .001 | 0.798           |  |
| Factor 5                | S     | 1.240    | 0.120 | 9.100  | < .001 | 0.000    | 0.507       | 1.750          |                      | 6         | 1.310    | 0.135 | 9.680  | < .001 | 0.796           |  |
|                         | 6     | 1.307    | 0.135 | 9.693  | < .001 | 0.795    | 0.686       | 1.759          |                      |           |          |       |        |        |                 |  |

Figure 2 Factor loading coefficient of confirmatory factor analysis. Notes: Estimate, parameter estimate value; SE, standard deviation; Stand. Estimate, standardized loading coefficient.

coefficient and McDonald's  $\omega$  coefficient of the second factor increased to 0.891 and 0.897; removing the remaining 18 items will decrease the two coefficients mentioned above. (Supplementary Figures 1–4)

## Discussion

The factor structure of the individual item of the translated version in this study is generally consistent with the threefactor structure of the original version, which suggests the universally applicable of JSE S-Version in both Chinese and English cultural contexts, and also demonstrates the accuracy of this translated version, providing preliminary evidence for the use of the translated version of the JSE S-Version in clinical teaching situations.

As with the original version of the Jefferson Scale, three factors were extracted by Principal component analysis in this study. Some researchers in past studies extracted more than three factors,<sup>23,24</sup> but these multi-factor models reduced the reliability of the scale, calling into question the applicability of the translated version of JSE they produced. In the author's opinion, the self-assessment of the JSE S-Version is for medical students, whereas after the direct translation of the JSE S-Version, the subject of each entry is a physician rather than a medical student, which means that medical students need to complete the self-assessed empathy scale as a third party, making it difficult for medical students to self-represent themselves in the role of a "doctor", which may result in the discrepancy of the results of the present study. In this study, after directly translating the original English version, we changed the subject from "doctor" to "doctor (me)" to strengthen the medical students' sense of self-substitution. Zhou Wenxia et al's study also showed that bringing students into the role of a doctor during the educational process can promote the construction of empathy.<sup>30,31</sup> The results of the study showed that the dimensions to which the individual item of our Chinese-translated version of the scale belonged were almost identical to those of the original English version, which suggests the value of the current Chinese-translated version as a valid translated version of the original English version of the JSE S-Version to be studied and used in mainland China.

In the current Chinese version, Exploratory factor analysis revealed that the factor loading of item 18 was lower than 0.3 in all dimensions, and the unique residual value was 0.867, which was much larger than 0.6, and after deleting the item 18, the cumulative variance explained rate of the three principal component factors could reach 53.928%. Confirmatory factor analysis further indicated standardized loading coefficient < 0.5 for item 18. Reliability analysis demonstrated increased internal consistency coefficient for dimension 2 after deleting item 18.32 These three lines of evidence collectively suggest that the item 18 has been shifted in the process of localization, and its meaning is relatively less clear. Similar characteristics were also shown in the previous studies on the reliability and validity of the JSE S-Version conducted in Japan, Iran and Turkey.<sup>15,17,33</sup> Item 18's poor performance may be attributed to its strongly negative contextual framing, which associates unfavorable outcomes or a burdensome atmosphere with the respondent. This affective context likely influenced students' judgment dimensions, resulting in a lower factor loading for Item 18. A similar phenomenon was observed in the French version of the JSE translated by Mariem in Tunisia, where items 1 and 18 also exhibited low factor loadings due to their negative wording.<sup>19</sup> Notably, other negatively phrased items did not demonstrate poor reliability or validity fit, suggesting that additional factors may have contributed to the suboptimal performance of Item 18. Although item 18 demonstrated a significantly lower standardized factor loading compared to other items, it exhibited acceptable item-total score correlation and favorable discrimination index effect sizes. This finding aligns with previous research by Hojat, who argued against excluding JSE items solely to improve latent variable model fit. Removing items may introduce incompatibility issues in comparative studies. Moreover, satisfactory item-total correlations and discrimination indices suggest that each item contributes meaningfully to the scale.<sup>29</sup> The structural instability of item 18, despite its acceptable item-total score correlation and satisfactory discrimination effects, presents an intriguing psychometric paradox. Future investigations should prioritize a thorough examination and potential refinement of this item's wording. Optimizing its linguistic formulation to enhance measurement properties represents a promising avenue for subsequent psychometric research. Additionally, in exploratory factor analysis, the unique residual values of item 15, 5, and 19 are slightly higher than 0.6, but the factor loadings in all dimensions are above 0.55, meeting the requirement of being greater than 0.4. Considering that the above three item have important professional significance in terms of their content and the subsequent confirmatory factor analysis and reliability analysis fitted them well, the item 15, 5, and 19 could be retained.

In this study, a Chinese translated version of the JSE S-Version was produced, and then the Chinese translated version was comprehensively evaluated from the perspectives of principal component analysis, exploratory factor analysis, confirmatory factor analysis, and reliability analysis. The results showed a high degree of factorial consistency with the original English version,<sup>34</sup> with satisfactory reliability and validity. These findings suggest that the Chinese translation of JSE S-Version shows promising potential as an assessment tool for measuring empathy levels among Chinese medical students, indicating its value for broader implementation and further research in the Chinese context.

This study also has some limitations. Although a reasonable sample size was calculated in advance, the sources of the sample were not sufficiently diverse to encompass populations from different cultures, regions, and economic backgrounds within China. The reliability and validity of the Chinese version of the JSE S-Version still require substantial sample sizes for further validation. While this study confirms that the Chinese version of the empathy scale demonstrates good internal consistency reliability, it lacks evaluations of other reliability measures, such as test-retest reliability and criterion-related validity. Future research should incorporate these aspects. Lastly, the assessment of empathy can generally be categorized into self-evaluation methods represented by scales and objective evaluations provided by patients or peers. Whether the empathy assessments conducted using the Chinese version of the JSE S-Version align well with objective evaluations still needs to be explored in future studies.

#### Conclusion

The Chinese-adapted JSE S-Version translated in this study represents a qualified psychometric instrument for measuring empathy. Overall, it demonstrates latent variable structure consistent with the original English version, with satisfactory reliability and validity outcomes. These findings support its preliminary application for evaluating empathy levels among medical students in Chinese cultural contexts. This study fills the gap in the research regarding the reliability and validity of the Chinese version of the JSE S-Version and confirms its applicability within Chinese linguistic and cultural environments.

#### Acknowledgments

All participants provided informed consent, in accordance with the Declaration of Helsinki.

#### Disclosure

The authors report no conflicts of interest in this work.

#### References

- 1. World Health Organization. Constitution of the World Health Organization. Am J Public Health Nations Health. 1946;36(11):1315–1323. doi:10.2105/AJPH.36.11.1315
- 2. Hojat M. Empathy in health professions, education, and patient care. 2016. doi:10.1007/978-3-319-27625-0
- 3. Herber-Valdez CR, Blow JA, Salazar TT, et al. The integrated curriculum and student empathy: a longitudinal multi-cohort analysis. *Adv Health Sci Educ.* 2024;29(4):1131–1153. doi:10.1007/s10459-023-10292-1
- 4. Hojat M, Gonnella JS, Mangione S, et al. Empathy in medical students as related to academic performance, clinical competence and gender. *Med Educ.* 2002;36(6):522–527. doi:10.1046/j.1365-2923.2002.01234.x
- 5. Quince T, Thiemann P, Benson J, Hyde S. Undergraduate medical students' empathy: current perspectives. *Adv Med Educ Pract*. 2016;7:443–455. doi:10.2147/AMEP.S76800
- Tamura H, Shikino K, Sogai D, et al. Association between physician empathy and difficult patient encounters: a cross-sectional study. J Gen Intern Med. 2023;38(8):1843–1847. doi:10.1007/s11606-022-07936-0
- 7. Berg K, Majdan JF, Berg D, Veloski J, Hojat M. Medical students' self-reported empathy and simulated patients' assessments of student empathy: an analysis by gender and ethnicity. *Acad Med.* 2011;86(8):984–988. doi:10.1097/ACM.0b013e3182224f1f
- O'Connor K, King R, Malone KM, Guerandel A. Clinical examiners, simulated patients, and student self-assessed empathy in medical students during a psychiatry objective structured clinical examination. Acad Psychiatry. 2014;38(4):451–457. doi:10.1007/s40596-014-0133-8
- Schweller M, Costa FO, Antônio M, Amaral EM, de Carvalho-Filho MA. The impact of simulated medical consultations on the empathy levels of students at one medical school. Acad Med. 2014;89(4):632–637. doi:10.1097/ACM.00000000000175
- 10. Berg K, Blatt B, Lopreiato J, et al. Standardized patient assessment of medical student empathy: ethnicity and gender effects in a multi-institutional study. *Acad Med.* 2015;90(1):105–111. doi:10.1097/ACM.0000000000529
- 11. Chen JY, Chin WY, Tsang JPY. How clinician examiners compare with simulated patients in assessing medical student empathy in a clinical exam setting. *Med Teach*. 2020;42(1):86–91. doi:10.1080/0142159X.2019.1665635
- 12. Al-Hashimi K, Said UN, Khan TN. Formative Objective Structured Clinical Examinations (OSCEs) as an assessment tool in UK undergraduate medical education: a review of its utility. *Cureus*. 2023;15(5):e38519. doi:10.7759/cureus.38519

- 13. Hojat M, Gonnella JS, Nasca TJ, Mangione S, Vergare M, Magee M. Physician empathy: definition, components, measurement, and relationship to gender and specialty. *Am J Psychiatry*. 2002;159(9):1563–1569. doi:10.1176/appi.ajp.159.9.1563
- 14. Hojat M. Empathy in patient care: antecedents, development, measurement, and outcomes. 2006.
- 15. Kataoka HU, Koide N, Ochi K, Hojat M, Gonnella JS. Measurement of empathy among Japanese medical students: psychometrics and score differences by gender and level of medical education. *Acad Med.* 2009;84(9):1192–1197. doi:10.1097/ACM.0b013e3181b180d4
- Paro HB, Daud-Gallotti RM, Tibério IC, Pinto RM, Martins MA. Brazilian version of the Jefferson scale of empathy: psychometric properties and factor analysis. BMC Med Educ. 2012;12:73. doi:10.1186/1472-6920-12-73
- 17. Shariat SV, Habibi M. Empathy in Iranian medical students: measurement model of the Jefferson scale of empathy. *Med Teach*. 2013;35(1):e913–918. doi:10.3109/0142159X.2012.714881
- Gerbase MW, Gustin MP, Bajwa N, Abbiati M, Baroffio A. Development and cross-national validation of a french version of the Jefferson scale of empathy for students. *Eval Health Professions*. 2022;45(3):288–296. doi:10.1177/01632787211033330
- Ghardallou M, Zedini C, Sahli J, Ajmi T, Khairi H, Mtiraoui A. Psychometric properties of a French version of the Jefferson scale of empathy. Int J Med Educ. 2022;13:205–214. doi:10.5116/ijme.62d2.8497
- Benmaamar S, Amri H, El Harch I, El Fakir S, El Rhazi K. Measurement properties of an Arabic version of the Jefferson Scale of Empathy (JSE) in Moroccan medical students. J Adv Med Educ Prof. 2024;12(1):37–44.
- Piumatti G, Costa MJ, Alcorta-Garza A, et al. Cross-national psychometric evaluation of the Jefferson scale of empathy-medical student version. Eval Health Professions. 2024;1632787241296540. doi:10.1177/01632787241296540
- 22. Voultsos P, Galanis P, Dafni MA, Velonaki VS, Andreou GN, Kovatsi L. The Greek Jefferson Scale of Empathy-Medical Student Version (JSE-S): psychometric properties and its associated factors. *Behav Sci.* 2024;14(3):195. doi:10.3390/bs14030195
- Šter MP, Šter B, Petek D, Gorup EC. Validation of Slovenian version of Jefferson scale of empathy for students. *Slovenian J Public Health*. 2014;53 (51):89–100. doi:10.2478/sjph-2014-0010
- Chen Z, Han Z, Dong Y, et al. Reliability and Validity Testing of the Jefferson Scale of Empathy in Chinese Medical Students for Measuring Their Empathy Ability. *Med Educ Res Pract*. 2022;30(6):723–726. doi:10.13555/j.cnki.c.m.e.2022.06.011
- Zhao J, Xiantao O, Li Q, et al. Role of narrative medicine-based education in cultivating empathy in residents. BMC Med Educ. 2023;23(1):124. doi:10.1186/s12909-023-04096-5
- 26. Li P, Weng L, Dong L. Empathy ability and influencing factors among pediatric residents in China: a mixed-methods study. *BMC Med Educ*. 2024;24(1):955. doi:10.1186/s12909-024-05858-5
- 27. Wang Q, Wang L, Shi M, et al. Empathy, burnout, life satisfaction, correlations and associated socio-demographic factors among Chinese undergraduate medical students: an exploratory cross-sectional study. *BMC Med Educ*. 2019;19(1):341. doi:10.1186/s12909-019-1788-3
- 28. Huang R, Zhou Z, Liu Y, et al. Empathy in undergraduate medical students: a multi-center cross-sectional study in China. *BMC Psychiatry*. 2024;24 (1):414. doi:10.1186/s12888-023-05350-2
- 29. Hojat M, LaNoue M. Exploration and confirmation of the latent variable structure of the Jefferson scale of empathy. *Int J Med Educ*. 2014;5:73–81. doi:10.5116/ijme.533f.0c41
- 30. Sarraf-Yazdi S, Teo YN, How AEH, et al. A scoping review of professional identity formation in undergraduate medical education. J Gen Intern Med. 2021;36(11):3511–3521. doi:10.1007/s11606-021-07024-9
- Zhou W, Wu J, Chen S, Wang J. To explore the application and practice of role play method in the cultivation of doctor-patient empathy ability of medical students. *China Higher Med Educ.* 2021;2021(12):23–24.
- 32. Rosellini AJ, Brown TA. Developing and Validating Clinical Questionnaires. Annu Rev Clin Psychol. 2021;17:55–81. doi:10.1146/annurev-clinpsy -081219-115343
- 33. Nazan Bilgel AO. Turkish version of the Jefferson scale of empathy psychometric properties. *Eur Sci J.* 2017;13:101. doi:10.19044/esj.2017. v13n20p101
- 34. Hojat M, Erdmann JB, Gonnella JS. Personality assessments and outcomes in medical education and the practice of medicine: AMEE Guide No. 79. Med Teach. 2013;35(7):e1267–1301. doi:10.3109/0142159X.2013.785654

Advances in Medical Education and Practice



Publish your work in this journal

Advances in Medical Education and Practice is an international, peer-reviewed, open access journal that aims to present and publish research on Medical Education covering medical, dental, nursing and allied health care professional education. The journal covers undergraduate education, postgraduate training and continuing medical education including emerging trends and innovative models linking education, research, and health care services. The manuscript management system is completely online and includes a very quick and fair peer-review system. Visit http://www.dovepress.com/testimonials.php to read real quotes from published authors.

Submit your manuscript here: http://www.dovepress.com/advances-in-medical-education-and-practice-journal

952 🖪 💥 in 🔼