



# Assessing measurement invariance of the SEGUE framework among Chinese and Ghanaian patients: A cross-sectional study

Jennifer Oppong<sup>a</sup>, Xinchun Liu<sup>a,\*</sup>, Qingyan Wang<sup>b</sup>, Chenhui Xie<sup>c</sup>, Yanjuan Lei<sup>d</sup>, Yuqiong Zhou<sup>e</sup>, Yan Peng<sup>f</sup>, Huiling Li<sup>g</sup>, Chengling Xing<sup>e</sup>

<sup>a</sup> Department of Clinical Psychology, Third Xiangya Hospital, Central South University, Changsha, Hunan, China

<sup>b</sup> Department of International Exchange and Cooperation, Third Xiangya Hospital, Central South University, Changsha, Hunan, China

<sup>c</sup> Department of Respiratory and Gastroenterology, People's Hospital of Xintian County, Yongzhou, Hunan, China

<sup>d</sup> Department of Nursing, Hospital of Traditional Chinese Medicine of Xintian County, Yongzhou, Hunan, China

<sup>e</sup> Department of Nephrology and Rheumatology, Third Xiangya Hospital, Central South University, Changsha, Hunan, China

<sup>f</sup> Department of Gastroenterology, Third Xiangya Hospital, Central South University, Changsha, Hunan, China

<sup>g</sup> Department of Pediatrics, Third Xiangya Hospital, Central South University, Changsha, Hunan, China

## ARTICLE INFO

### Keywords:

SEGUE framework

Measurement invariance

Doctor patient communication skills

## ABSTRACT

**Objective:** Measurement invariance assesses the psychometric equivalence of a construct across groups or across time. This study aimed to find out whether the newly modified 5-point Likert scale SEGUE Framework for evaluating doctor's communication skills demonstrates measurement invariance among Chinese and Ghanaian patients.

**Methods:** Measurement invariance tests were conducted using two samples collected from the two cultures respectively (China:  $N = 598$  and Ghana:  $N = 591$ ). Confirmatory Factor Analysis (CFA) was used to test for the factor structure. Multiple-group CFA was used to test the measurement invariance of the SEGUE between the two samples.

**Results:** The same five-factor model showed acceptable model fit in both Chinese and Ghanaian samples. Cross-cultural measurement invariance tests revealed that scalar invariance was observed across the two samples.

**Conclusion:** Cross-cultural adaptations of instruments should exhibit measurement invariance to ensure that measures are interpreted and used consistently. Findings of our study suggest that the SEGUE Framework is invariant across the two cultures, which increases confidence in data collection and comparison.

**Innovation:** This is the first study to establish measurement invariance among Chinese and Ghanaians using the SEGUE Framework which assesses doctors' communication from patients' perspectives.

## 1. Introduction

Doctor-patient communication refers to the process by which doctors and patients develop understanding, reach consensus, and establish trust and cooperation. This can be done through effective information exchange by focusing on disease prevention, diagnosis, treatment, rehabilitation, and other related issues. The capacity to acquire information in order to enable a correct diagnosis, counsel properly, give therapeutic guidance, and maintain caring relationships with patients are all part of a doctor's communication and interpersonal skills [1]. Furthermore, these communication skills are the fundamental clinical skills required for the effective delivery of health care, with the ultimate goal of obtaining expected health outcomes and patient satisfaction [2].

Instruments or tools that are meant to be used reliably in numerous contexts, such as checklists, frameworks, guidelines, or models, can provide a

structural foundation for attempts to teach, measure, study, and enhance the communication skills of both physicians-in-training and physicians-in-practice. There exist a number of instruments and tools used in assessing communication skills of physicians, such as the SEGUE Framework, the Four Habits Coding Scheme, the Communication Assessment Tool (CAT), the Liverpool Communication Skills Assessment Scale, the Global consultation Rating Scale, the Roter Interaction Analysis System, the Calgary Cambridge Observation Guide Checklist [3], etc. Among these assessment tools, the SEGUE Framework is a notable one that has been widely used and referred to. The SEGUE Framework is a checklist of medical communication tasks intended to aid in the teaching and assessment of communication skills, as well as doctor-patient communication research [4]. A number of studies [4-8] conducted over the past 10 years have demonstrated that the SEGUE Framework has good reliability and validity and may be used to accurately assess doctors' communication skills. Although the SEGUE

\* Corresponding author.

E-mail address: [angelaliu23@sina.com](mailto:angelaliu23@sina.com) (X. Liu).

Framework has evidence of its usage in teaching, research, and assessments [4] and has been widely used among different population [5,9,10], research on its measurement invariance is still lacking.

Measurement invariance assesses the (psychometric) equivalence of a construct across groups or measurement occasions and demonstrates that a construct has the same meaning to those groups or across repeated measurements [11]. It applies equally to research topics in clinical, cognitive, social, experimental, cross-cultural, and developmental psychology [11]. Measurement equivalence is another term for measurement invariance which is used interchangeably in the literature [12,13], but for the purpose of clarity, we used measurement invariance in this study. Based on the fact that words and construct meanings (and their translated equivalents) may differ in different languages and cultural contexts, it is recommended to test measurement invariance for each group under study [14]. Cross-cultural studies are systematic comparisons of different cultures with the goal of understanding how cultural context influences human behavior [15]. To our knowledge, there is no study on measurement invariance of the SEGUE Framework between China and Ghana; therefore, establishing a cross-cultural invariance will be of great importance in this field of study, thereby filling the gap in the literature. Owing to the fact that both Ghana [16] and China [17] are collectivist cultures, whereby the wants and objectives of the group as a whole are prioritized over the needs and preferences of each individual. In such societies, each person's identity is fundamentally shaped by their connections to other individuals and their relationships with other group members, as opposed to individualistic cultures. Cultural differences may influence the way these two cultures perceive communication between physicians and patients.

Ghana was chosen to be compared to China because the first author is an international student from Ghana and is also among the English-speaking population in Ghana, so with the approval of her mentor (the principle investigator of this study/the corresponding author), the team decided to choose Ghana as a comparator for the purpose of exploring the cross-cultural differences of the SEGUE Framework and for the convenience of acquiring samples as well. Considering the examination of such potential cultural differences, researchers have convincingly argued that prior to studying the cross-cultural differences in the levels and correlates associated with the measures, the cross-cultural equivalence of psychological measures of interest should be examined [18]. To regard the scale as being consistent across populations and to make legitimate and understandable comparisons of the variations in scores, empirical evidence of measurement invariance is required. Specifically, scalar-level invariance must be achieved for mean comparisons between the groups [19]. The idea of measurement invariance states that the same measurement tool measures the same construct in the same way across cultures, or the same measurement instrument can be used to measure the same construct in different cultures [20]. The terms culture and country are used synonymously in the text, which is significant given that the samples are each defined by a set of common languages, shared beliefs, values, and social norms. It is common to find literature on instrument adaptations across nations that uses the word "culture" in such a flexible manner [21]. Guo & Wang [5] modified the SEGUE Framework on a 5-point Likert scale for both patients and doctors' versions in 2021. This makes assessment more precise compared to the original "yes" or "no" checklist. We think patients' perspectives are very important for effective or high-quality doctor-patient communication; hence, we decided to conduct a measurement invariance study on the newly modified 5-point Likert scale SEGUE Framework (patients' version).

Therefore, this paper aimed to determine whether the SEGUE Framework has measurement invariance and can be used across different cultures and populations, specifically among Ghanaian and Chinese patients. Based on reviewing the literature, it was expected that both configural and metric invariances would hold but not a full scalar invariance. Again, the factor structures of the SEGUE Framework for the two samples should be similar. The two most popular methods for testing measurement invariance of cross-cultural data are methods for detecting differential item functioning (DIF) [22], developed in the context of item response theory (IRT) [23], and Multiple-group CFA [24]. Recently, a less well-known but interesting

method has been discovered that combines multiple-group latent class analysis with latent class factor analysis. The three methods all have the same basic idea, which is to define a measurement model by comparing the latent structure for several groups within a single model [20]. For Multiple-group CFA, we fit a measurement model for each group separately, then compare the fit of many nested models that are formed by constraining the parameter estimates between the groups [25]. There are four types (levels) of measurement invariance that are assessed using Multiple-group CFA: configural (equal dimensional structure), metric (equal factor loadings), scalar (equal factor loadings and intercepts or thresholds), and strict factorial invariance (equal factor loadings, intercepts or thresholds, and indicator residual variances) [25]. Most scholars suggest that achieving scalar invariance is required to do a mean comparison across different populations, even though additional levels that are stricter could be tested for [26,27].

## 2. Methods

### 2.1. Participants and procedure

Participants were recruited between March and November 2022. In Ghana, nearly half of the participants were recruited from one teaching hospital. It is a referral hospital that receives patients from all parts of the country, and it's also a large teaching hospital. The rest came from four district hospitals under the Ghana Health Service in the southern part of Ghana. These were selected towns located in the Ashanti Region of Ghana. Hospitals under this jurisdiction were mainly primary hospitals that act as a three-tier system where health services are provided at district, sub-district, and community levels. In order to meet the inclusion criteria, trained research assistants (who were also nurses) in collaboration with some staff working in the mentioned hospitals identified potential participants through patients' files. Patients' files were selected based on their last visit to the hospital. This was done after getting ethical approval from these hospitals. Research assistants screened the patients' files to identify patients from the ages of 18 and above, as well as checked for other information such as patients' occupations to give a clue as to whether they are likely to understand English or not. The identified patients were then contacted through telephone calls. Those participants who did not meet the criteria (could not speak English) were exempted from the study, and those deemed eligible for the study were explained the study purpose and asked if they agreed or not to participate. The time and dates were scheduled depending on their availability for the interview. For those who had a face-to-face interview, interviews were conducted in the outpatient department of the respective hospital. They were given an information sheet that entails the study aims and other information about the study to read before signing the written consent form. The process took a maximum of 15 min for each patient. Those who could not make the face-to-face interview, on the other hand, answered the questionnaire on the telephone after consent was sought, signed, and sent through an online medium. The signed forms were then sent to the research assistants via the same online platform prior to answering the questionnaire over the telephone.

With the sample from China, these were Chinese nationals, particularly those drawn from the Hunan Province. Considering the similarity of samples from the two countries, we selected one big teaching hospital and three county hospitals where the economic status is far below the average level in China to be our study settings. They were selected using a convenience sampling method, whereby patients who were present at the time the study began could participate. Patients who visited the hospital were asked if they were willing to participate. Participants answered the paper questionnaire.

### 2.2. Measures

The original SEGUE Framework was developed and validated by Makoul [4], which has a yes/no response, and was revised in China in 2021 [5], which is rated on a 5-point Likert scale. The scale has gained

considerable acceptance in teaching, assessment, and research areas. The framework has shown generalizability throughout the health professions, from pediatrics to gynecology and oncology to psychology, in addition to its application at several medical schools. There are five dimensions and 25 items in the SEGUE Framework for both the English and Chinese versions. There are 5 items in the “set the stage” dimension, 10 items in the “elicit information” dimension, 4 items in the “give information” dimension, 4 items in the “understand the patient's perspective” dimension, and 2 items in the “end the encounter” dimension. It was rated on a 5-point Likert scale ranging from 1 to 5 (1 = never, 2 = occasionally, 3 = sometimes, 4 = mostly, 5 = all the time). This makes responses relatively easy to compute and makes answering questions simpler for the respondent because it does not compel the participant to give a clear-cut “yes” or “no” response, forcing them to take a position on a subject instead of allowing them to express some level of agreement. Additionally, the responses take into account participants' neutral responses. The Ghanaian samples that met the inclusion criteria answered the English questionnaire, while the Chinese answered the Chinese version of the questionnaire. “Set the stage” describes the stage of the patient's opening statement about his or her concerns; “Elicit information” describes the stage of eliciting and exploring the patient's full set of problems, including the causes and effects of the problems; “Give information” refers to the stage in which a doctor discloses details about a patient's health; “understand the patient's perspective” refers to the stage of the doctor's understanding of the patient's efforts and concerns and establishing relationships with their patients, which may run through the entire communication process. At the “end of the encounter” stage, the doctor reviews the next steps with the patient and inquires if there is anything else they would like to talk about [8].

### 2.3. Ethical consideration

Ethical approval of the research was obtained from the Third Xiangya Hospital of Central South University. Additionally, permission to start the study was obtained from all the hospitals where participants were selected following approval from their institutional review board. All participants were informed about the purpose of the study. Written informed consent was obtained from participants before answering the questionnaire. Participation was voluntary, and they were free to withdraw if they felt their rights as participants were being violated. To ensure participants' privacy, they were guaranteed confidentiality and anonymity throughout the study.

### 2.4. Analyses

Internal consistencies and descriptive statistics were conducted using SPSS version 25. The most widely used measure of internal consistency reliability is Cronbach alpha ( $\alpha$ ) [28]. A reliability coefficient (Cronbach alpha) of 0.70 or higher is considered acceptable [28]. A Multiple-group CFA was conducted using MPlus 8.3. Data was analyzed in two steps. CFAs were conducted for the two samples separately to examine the factor structure for each country. Maximum likelihood estimations were used. The following indices were used to assess the fit of the model to the data: the chi-square ( $\chi^2$ ) statistic and the related degrees of freedom (df), the  $\chi^2$ /df ratio, the Standardized Root Mean Square Residual (SRMR), the Tucker-Lewis Index (TLI), the Comparative Fit Index (CFI), and the Root Mean Square Error of Approximation (RMSEA) were used to evaluate the goodness of fit of the models. Criteria for the model fit indices ranged from  $\chi^2$ /df  $\leq 3$ , TLI and CFI values  $\geq 0.9$ . Furthermore, SRMR and RMSEA values  $\leq 0.08$  were considered reasonable, but values  $\leq 0.05$  indicate excellent fit [29]. Subsequently, Multiple-group CFA were conducted to examine whether the SEGUE demonstrated measurement invariance across Chinese and Ghanaian patients. A reliability test was conducted by collecting data from 20 to 30 participants who were not part of the sample. In this study, Cronbach alpha was 0.90 and 0.93 in China and Ghana, respectively. Measurement invariance was conducted in an orderly manner where an initial analysis (Model 0), with the only requirement for invariance being that the same parameters existed for both groups (configural

invariance). With the configural invariance the same model that was used and tested separately was then tested simultaneously across the two countries without imposing any constraints on the model's parameter. Subsequently, analyses added the additional invariance constraints of equivalent factor loadings to test for metric invariance after configural invariance was established (Model 1). We placed restrictions on factor loadings and item intercepts to test for scalar invariance once metric invariance was achieved (Model 2). Three levels of measurement invariance including configural (equal factor structure), metric (equal factor loadings), and scalar (equal factor loadings and intercepts/thresholds) [13] were assessed. The criteria for the Measurement invariance testing was  $\Delta$ RMSEA and  $\Delta$ CFI. Cheung and Rensvold [30], stated that an insignificant difference in the  $\Delta$ CFI  $\leq 0.010$  and  $\Delta$ RMSEA  $\leq 0.015$  for the difference between the metric and configural, scalar and metric was evidence of invariance.

## 3. Results

### 3.1. Descriptive statistics of the two samples

The samples comprised 1189 patients, out of which 598(50.3%) came from China and 591(49.7%) from Ghana. Their ages ranged from 18 to 69 years (see Table 1 for details).

### 3.2. Fit parameters for SEGUE of the two samples

The model fit of the CFA models for both samples are shown in (Table 2) below. The values for  $\chi^2$ /df were  $< 3$ , RMSEA and SRMR were both  $< 0.08$ , CFI and TLI values were all  $> 0.90$ . Specifically,  $\chi^2$ /df = 2.067, RMSEA = 0.042, CFI = 0.916, TLI = 0.904, SRMR = 0.044 for the Ghanaian sample, and  $\chi^2$ /df = 2.356, RMSEA = 0.048, CFI = 0.959, TLI = 0.954, SRMR = 0.029 for the Chinese sample. Results indicated that the model fit of the five-factor model for both samples is acceptable.

The five-factor model and its structure among Ghanaian sample is shown in Fig. 1 and that of Chinese sample is shown in Fig. 2.

Fig. 1 presents the factor structure of the five-factor CFA for the Ghanaian sample ( $n = 591$ ); goodness of fit indices:  $\chi^2$  value = 541.425, df = 262, RMSEA = 0.042, SRMR = 0.044, CFI = 0.916, TLI = 0.904.

**Table 1**  
Descriptive statistics of the study samples ( $N = 1189$ ).

| Variables | N   | %    |
|-----------|-----|------|
| China     |     |      |
| Gender    |     |      |
| Male      | 277 | 46.3 |
| Female    | 321 | 53.7 |
| Age       |     |      |
| $\leq 30$ | 133 | 22.2 |
| 31–40     | 113 | 18.9 |
| 41–50     | 110 | 18.4 |
| $> 50$    | 242 | 40.5 |
| Education |     |      |
| Primary   | 237 | 39.6 |
| Junior    | 207 | 34.6 |
| High      | 131 | 21.9 |
| Tertiary  | 23  | 3.8  |
| Ghana     |     |      |
| Gender    |     |      |
| Male      | 291 | 49.2 |
| Female    | 300 | 50.8 |
| Age       |     |      |
| $\leq 30$ | 306 | 51.8 |
| 31–40     | 169 | 21.6 |
| 41–50     | 87  | 14.7 |
| $> 50$    | 29  | 4.9  |
| Education |     |      |
| Middle    | 139 | 23.5 |
| High      | 216 | 36.5 |
| Tertiary  | 236 | 39.9 |

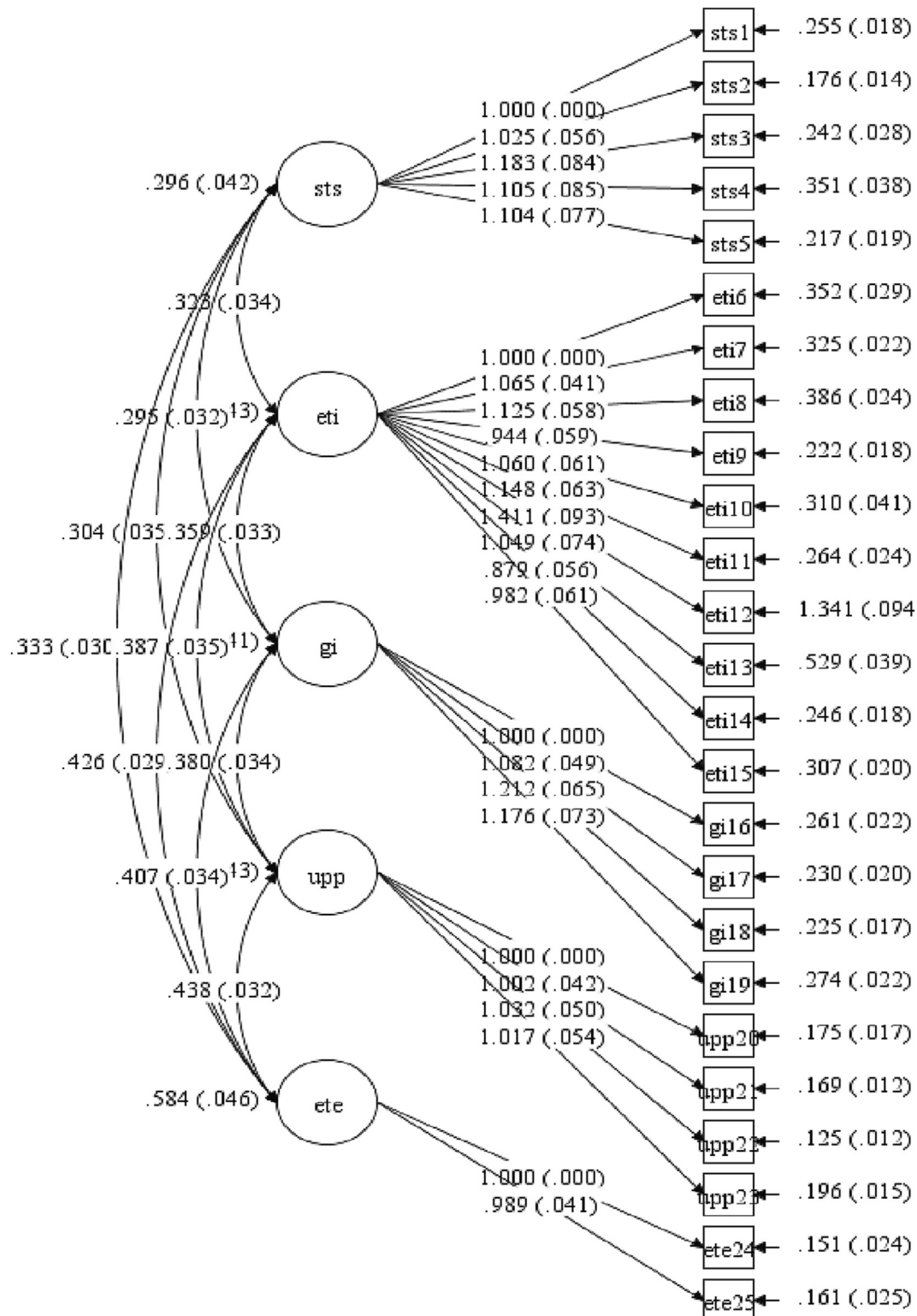
Note: N = sample size; % = percentage.

**Table 2**

Fit parameters for SEGUE of the two samples.

|           | $\chi^2$ | df  | $\chi^2/df$ | RMSFA | CFI   | TLI   | SRMR  |
|-----------|----------|-----|-------------|-------|-------|-------|-------|
| Threshold |          |     | <3          | <0.08 | >0.9  | >0.9  | <0.08 |
| Ghanaian  | 541.425  | 262 | 2.067       | 0.042 | 0.916 | 0.904 | 0.044 |
| Chinese   | 622.012  | 264 | 2.356       | 0.048 | 0.959 | 0.954 | 0.029 |

Note:  $\chi^2$  = chi-square; df = degrees of freedom; RMSEA = Root Mean Square Error of Approximation; SRMR = Standardized Root Mean Square Residual; TLI = Tucker-Lewis Index; CFI = Comparative Fit Index.

**Fig. 1.** The five-factor model and its structure among Ghanaian sample.



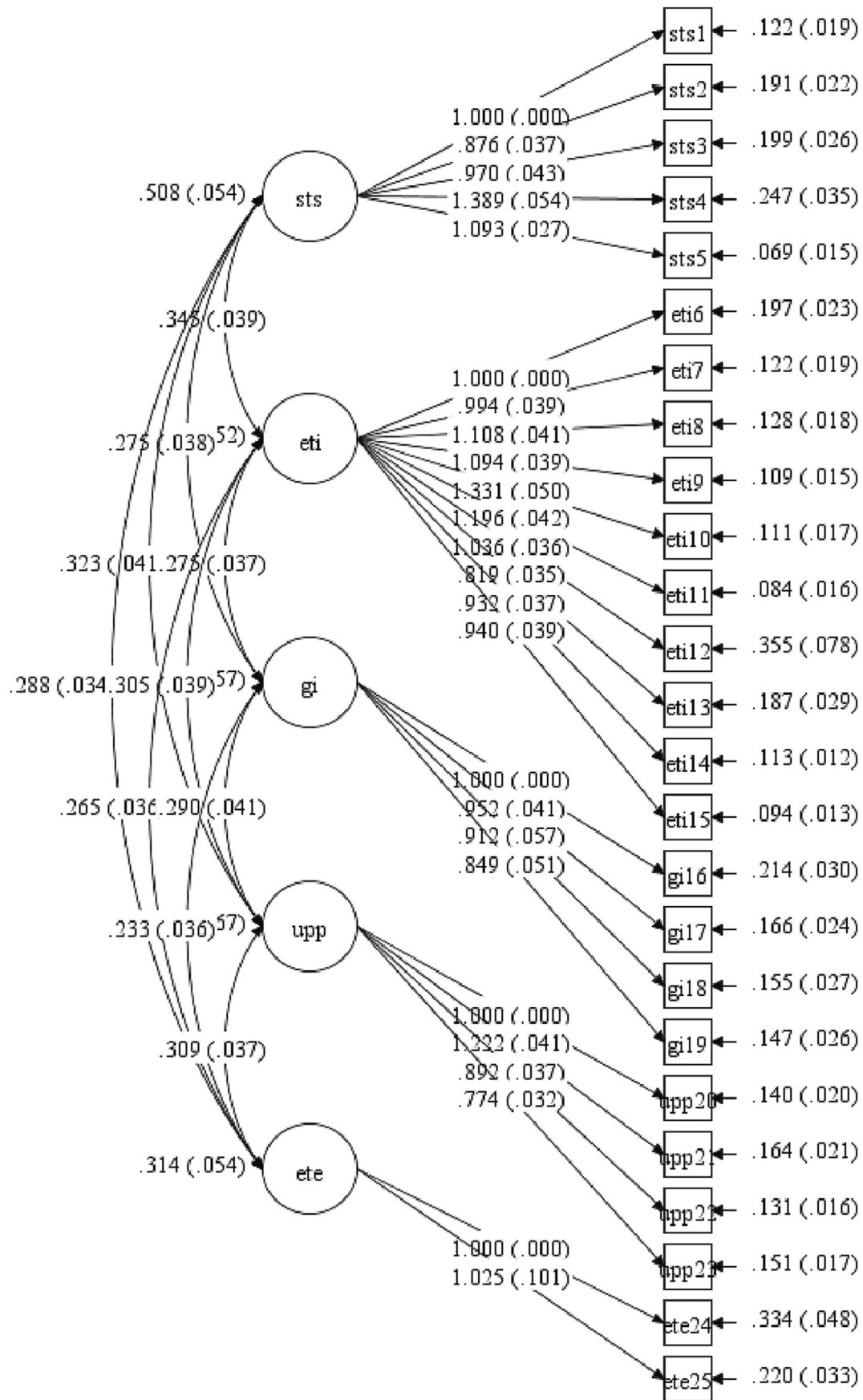


Fig. 2. The five-factor model and its structure among Chinese sample.

sts = set the stage; (sts1, sts2, sts3, sts4, sts5), eti = elicit the information; (eti6, eti7, eti8, eti9, eti10, eti11, eti12, eti13, eti14, eti15), gi = give information; (gi16, gi17, gi18, gi19), upp = understand the patient's perspective; (upp20, upp21, upp22, upp23), ete = end the encounter; (ete24, ete25) CFA = confirmatory factor analysis; ( $\chi^2$ ) the chi-square statistic; df = its related degrees of freedom, SRMR = Standardized Root Mean Square Residual, TLI = Tucker-Lewis Index, CFI = Comparative Fit Index, RMSEA = Root Mean Square Error of Approximation.

Fig. 2 presents the factor structure of the five-factor CFA for the Chinese sample ( $n = 598$ ); goodness of fit indices:  $\chi^2$  value = 622.012, df = 264, RMSEA = 0.048, SRMR = 0.029, CFI = 0.959, TLI = 0.954.

sts = set the stage; (sts1, sts2, sts3, sts4, sts5), eti = elicit the information; (eti6, eti7, eti8, eti9, eti10, eti11, eti12, eti13, eti14, eti15), gi = give information; (gi16, gi17, gi18, gi19), upp = understand the patient's perspective; (upp20, upp21, upp22, upp23), ete = end the encounter; (ete24, ete25).

CFA = confirmatory factor analysis; ( $\chi^2$ ) the chi-square statistic; df = its related degrees of freedom, SRMR = Standardized Root Mean Square Residual, TLI = Tucker-Lewis Index, CFI = Comparative Fit Index, RMSEA = Root Mean Square Error of Approximation.

### 3.3. Measurement invariance across the two countries

Table 3 presents the Multiple-group CFA summary fit indices for the three nested models.

- 1) The difference of CFI between the Metric model and the Configural model is 0.000, <0.010, and the difference of RMSEA is 0.000, <0.015, indicating that the equivalence on factor loadings between the two groups held and Metric level invariance was achieved.
- 2) The difference of CFI between Scalar model and Metric model is 0.010, equal to 0.010, and the difference of RMSEA is -0.002, <0.015. This indicates that the equivalence on both factor loadings and intercepts between the two groups held and scalar level invariance was achieved.

Changes in fit indicated that the metric and configural model did not differ significantly (i.e.  $\Delta$  CFI and  $\Delta$  RMSEA = 0.000) which is <0.010 and < 0.015 respectively [30]. Changes in scalar and metric did not significantly differ (i.e.  $\Delta$  CFI = 0.010 and  $\Delta$  RMSEA -0.002 which is  $\leq 0.010 < 0.015$  [30]. Therefore, it shows that the two scales are equivalent between Chinese and Ghanaians.

## 4. Discussion and conclusion

### 4.1. Discussion

The SEGUE Framework is a checklist of medical communication tasks intended to aid in the teaching and assessment of communication skills, as well as doctor-patient communication research [4]. The main objective of the study was to determine the measurement invariance of the SEGUE Framework among Chinese and Ghanaian patient samples. The significance of measurement invariance in cross-cultural comparisons is receiving more and more attention [31]. Our study is the first to examine the measurement invariance of the SEGUE Framework rated on a 5-point Likert scale, specifically between Ghanaian and Chinese cultures. Descriptive statistics, internal consistencies, and CFAs were calculated prior to testing the scale's measurement invariance across the two countries. The

25-item SEGUE appears to accurately assess the five dimensions of "set the stage," "elicit information," "give information," "understand the patient's perspective," and "end the encounter," according to results with excellent internal consistency coefficients and its consistency with previous studies [5,10]. The CFA for the factor structure shows adequate support for the five-factor model in the two samples. Both the English and Chinese versions of SEGUE were equivalent on all five domains, thereby satisfying our first hypothesis. An unconstrained model with good fit would suggest that groups share a common factor structure [13]. We also tested for measurement invariance between the two countries. Three models, including configural, metric, and scalar were tested, and none of their fits between the two groups significantly decreased. Thus, the differences between the models were non-significant and may suggest that the overall structure of the scale and responses are valid equally across the two groups [32]. Configural invariance determines whether or not the same items accurately measure our construct across the two populations. Achieving invariance of factor loadings indicates that participants interpret the construct similarly across administrations. In other words, in each group, the same latent factor was being measured. Achieving scalar or strong invariance establishes the justification for mean comparisons across time or between groups [26].

Establishing scalar invariance allows multi-group comparisons of factor means (like  $t$ -tests or ANOVA) [27]. By so doing, it is certain that any statistically significant differences in group means are not caused by variations or differences in scale properties over time. For instance, if we compare the means of the observed scores on the scale for two cultural groups and discover a significant difference, our ideal claim would be that the difference is the result of real differences on the construct we are trying to measure [21]. Thus, our results show that the SEGUE is an appropriate measure to investigate physicians' communication skills between Chinese and Ghanaians. Here, too, the second hypothesis was accepted. This means that differences in latent communication scores between Chinese and Ghanaians can be interpreted as meaningful and not as a result of measurement error or variations in measurement structure between the groups. This may mean that responses to the scale and its general structure may be equally valid for participants within these two groups. Clear evidence of measurement invariance on the SEGUE suggests that even though China and Ghana have sociocultural differences, the patients understood the questionnaire in a similar way. The fact that measurement invariance was established even with different cultures may be partially due to the clear understanding of the items by each of the groups. As a result, while health, communication, and other topics may be viewed differently by different cultures [33], SEGUE measures physicians' communication skills, and the scale's meaning is similar for participants from China and Ghana. Even though additional steps can provide an even stricter test of measurement invariance, researchers generally agree that evaluating configural, metric, and scalar invariance is sufficient for establishing measurement invariance [26,27]. There is evidence of full scalar invariance, which is consistent with previous studies [34], as opposed to some researchers who argue that the achievement of full scalar invariance is rarely established, especially in diverse cultural contexts [35,36]. In some studies, full scalar invariance was not achieved but partial scalar invariance [13,37-39].

This is the first study to compare the measurement invariance of the SEGUE between Chinese and Ghanaians, and the results demonstrated an overall acceptable measurement invariance for the SEGUE. As we used the newly adapted 5-point Likert scale SEGUE Framework, whose evaluation we think would be more explicit than the original version with "yes" or "no" assessments, this might become a good strength of our study. This means the 25 items could be used for cross-cultural comparison at a more precise level. The results of this study should be used as a guide for future studies in this area of research. Furthermore, the 5-point Likert scale SEGUE Framework is designed to be applied among patients in order to explore their perspectives on their doctor's communication skills; this might become another strength of our study as there is no research of this kind regarding SEGUE. However, the fact that the samples were limited to patients

**Table 3**  
Measurement invariance across the two countries.

|                | $\chi^2$ | Df  | CFI   | RMSEA | Model comparison | $\Delta$ CFI | $\Delta$ RMSEA |
|----------------|----------|-----|-------|-------|------------------|--------------|----------------|
| M0(Configural) | 1832.478 | 176 | 0.917 | 0.065 |                  |              |                |
| M1(Metric)     | 1832.478 | 176 | 0.917 | 0.065 | M1 VS M0         | 0.000        | 0.000          |
| M2(Scalar)     | 2008.548 | 156 | 0.907 | 0.067 | M2 VS M1         | 0.010        | -0.002         |

could also be a weakness of the study because the findings could not be generalized to other populations in these two countries. Participants in the study were recruited by using convenience sampling, which was restricted to certain areas and medical facilities. The generalization of the study was also limited. Therefore, future research should include participants from different provinces or regions in the two countries. Additional measurement invariances on demographic variables such as educational levels and gender could be explored in future studies. Since samples from China and Ghana were used in the current study, the measurement invariance of the scale on samples from America, Australia, Italy, and England will need to be investigated in future studies. The results show that the SEGUE scores can be meaningfully compared across countries, but with caution.

#### 4.2. Innovation

A crucial and more well-known methodological aspect of comparative and developmental psychological science is measurement invariance. Establishing cross-cultural measurement invariance is very important but often overlooked in studies that conduct comparative research as well as psychological studies. This is the first study to measure doctors' communication skills between Chinese and Ghanaians, specifically testing the cross-cultural measurement invariance among these two cultures. Although the SEGUE Framework has been used in different populations, there has not been any explicit study to show that the SEGUE Framework has the same meaning in the two cultures. The results of this study suggest that the SEGUE Framework is a reliable tool for medical encounters in physician-patient communication in the two cultures. Prior to this study, research on the measurement invariance of the SEGUE was heavily underestimated, and thus the scale could not be used in different contexts.

This study proposed that patients in these two cultures might be able to assess and report doctors' communication skills based on their own (patients') perceptions. The Ghanaian and Chinese versions of SEGUE are invariant, which allows for comparison of means across different cultures, which supports its validity. The results of CFA suggest a five-factor structure: one factor loaded on each of the five dimensions (set the stage, elicit the information, give the information, understand the patient's perspective, and end the encounter) across the two cultures. Some studies have argued that full cross-cultural invariance on a measurement scale is unattainable, but in this study, there was an achievement of full scalar invariance on all items of the SEGUE in the two countries.

#### Appendix A

| The SEGUE Framework (English version)                                                                              |   |   |   |   |   |
|--------------------------------------------------------------------------------------------------------------------|---|---|---|---|---|
| 1 = never; 2 = occasionally; 3 = sometimes; 4 = mostly; 5 = all the time                                           |   |   |   |   |   |
| <b>Set the stage</b>                                                                                               |   |   |   |   |   |
| 1. greet me appropriately                                                                                          | 1 | 2 | 3 | 4 | 5 |
| 2. explain the reason for the consultation                                                                         | 1 | 2 | 3 | 4 | 5 |
| 3. Introduce the outline agenda to me for visit (e.g., issues, sequence)                                           | 1 | 2 | 3 | 4 | 5 |
| 4. Make a personal connection with me during visit (e.g., go beyond medical issues at hand)                        | 1 | 2 | 3 | 4 | 5 |
| 5. Maintains my privacy (e.g., close door)                                                                         | 1 | 2 | 3 | 4 | 5 |
| <b>Elicit the information</b>                                                                                      |   |   |   |   |   |
| 6. Elicit my view of health problem and/or progress                                                                | 1 | 2 | 3 | 4 | 5 |
| 7. Explore physical/physiological factors of my medical issues                                                     | 1 | 2 | 3 | 4 | 5 |
| 8. Explore psychosocial/emotional factors of my medical issues (e.g., living situation, family, relations, stress) | 1 | 2 | 3 | 4 | 5 |
| 9. Discuss antecedent treatments with me (e.g., self-care, last visit, other medical care)                         | 1 | 2 | 3 | 4 | 5 |
| 10. Discuss how health problem affects my life (e.g., quality-of life)                                             | 1 | 2 | 3 | 4 | 5 |
| 11. Discuss lifestyle issues/prevention strategies with me (e.g., health risks)                                    | 1 | 2 | 3 | 4 | 5 |
| 12. Avoid asking me directive/leading questions                                                                    | 1 | 2 | 3 | 4 | 5 |
| 13. Give me opportunity/time to talk (e.g. don't interrupt)                                                        | 1 | 2 | 3 | 4 | 5 |
| 14. Listen. Give me undivided attention (e.g., face me, verbal acknowledgement, negative feedback)                 | 1 | 2 | 3 | 4 | 5 |
| 15. Check/clarify information with me (e.g., recap, ask "how much")                                                | 1 | 2 | 3 | 4 | 5 |
| <b>Give information</b>                                                                                            |   |   |   |   |   |
| 16. Explain rationale for diagnostic procedures for me (e.g., exam, tests)                                         | 1 | 2 | 3 | 4 | 5 |
| 17. Teach me about my own body & situation (e.g., provide feedback from exam/tests, explain anatomy/diagnosis)     | 1 | 2 | 3 | 4 | 5 |
| 18. Encourage me to ask questions/check understanding                                                              | 1 | 2 | 3 | 4 | 5 |

(continued on next page)

#### 4.3. Conclusion

Cross-cultural adaptations of instruments should exhibit measurement invariance to ensure that measures are interpreted and used consistently. The evidence of invariance of the SEGUE Framework demonstrates that the SEGUE Framework is a valid instrument for investigating cross-cultural differences in physicians' communication skills. Psychological research relies heavily on testing for measurement invariance to make sure that comparisons between different groups are valid and meaningful. Without tests of measurement invariance, "we cannot assume the same construct is being assessed across groups by the same measure" (p. 108) [40]. Therefore, regardless of the subject or discipline, measurement invariance testing is a crucial addition to our toolbox of statistical techniques that aids in enhancing the reliability and validity of our research. There is currently enough data to confidently conduct additional empirical validity research using the SEGUE Framework in the two contexts.

#### Funding

The study was not funded.

#### Credit authorship contribution statement

JO: Conception, design, data collection and analysis, drafted and edit the manuscript. XL: Conception, design, critically reviewing the manuscript. QW, CX, YL, YZ, YP, HL and CX: Data collection and analysis. All authors gave their final approval and agreed to be accountable for all aspects of the work.

#### Declaration of Competing Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

#### Acknowledgement

We acknowledge all the hospitals that agreed to provide us the samples. We also thank those colleagues who helped us on distributing and collecting the questionnaires.

(continued)

| The SEGUE Framework (English version)                                    |   |   |   |   |   |
|--------------------------------------------------------------------------|---|---|---|---|---|
| 1 = never; 2 = occasionally; 3 = sometimes; 4 = mostly; 5 = all the time |   |   |   |   |   |
| 19. Adapt to my level of understanding (e.g., avoid/explain jargon)      | 1 | 2 | 3 | 4 | 5 |
| <b>Understand the Patient's perspective</b>                              |   |   |   |   |   |
| 20. Acknowledge my accomplishments/progress/challenges                   | 1 | 2 | 3 | 4 | 5 |
| 21. Acknowledge waiting time                                             | 1 | 2 | 3 | 4 | 5 |
| 22. Express caring, concern, empathy for me                              | 1 | 2 | 3 | 4 | 5 |
| 23. Maintain a respectful tone                                           | 1 | 2 | 3 | 4 | 5 |
| <b>End the encounter</b>                                                 |   |   |   |   |   |
| 24. Ask if there is anything else I would like to discuss                | 1 | 2 | 3 | 4 | 5 |
| 25. Review next steps with me                                            | 1 | 2 | 3 | 4 | 5 |

© Copyright 1993/2000 Gregory Thomas Makoul- All Rights Reserved.

Appendix B

| The SEGUE Framework (Chinese version)           |   |   |   |   |   |
|-------------------------------------------------|---|---|---|---|---|
| 1 = 从不; 2 = 偶尔; 3 = 有时; 4 = 大多数时候; 5 = 一直 准备    |   |   |   |   |   |
| 1. 有礼貌地称呼病人                                     | 1 | 2 | 3 | 4 | 5 |
| 2. 说明此次问诊的理由(了解情况/进一步诊断治疗/汇报上级医师)               | 1 | 2 | 3 | 4 | 5 |
| 3. 介绍问诊和查体的过程(如问诊的内容、先后顺序等)                     | 1 | 2 | 3 | 4 | 5 |
| 4. 建立个人信任关系(如适当的自我介绍/讨论一些目前疾病以外的话题)             | 1 | 2 | 3 | 4 | 5 |
| 5. 保护病人的隐私(如关门等)/尊重病人的选择权/隐私权                   | 1 | 2 | 3 | 4 | 5 |
| <b>信息收集</b>                                     |   |   |   |   |   |
| 6. 保护病人的隐私(如关门等)/尊重病人的选择权/隐私权                   | 1 | 2 | 3 | 4 | 5 |
| 7. 系统询问影响疾病的物理/生理因素                             | 1 | 2 | 3 | 4 | 5 |
| 8. 系统询问影响疾病的社会、心理/情感因素(如生活水平、社会关系、生活压力等)        | 1 | 2 | 3 | 4 | 5 |
| 9. 与病人讨论既往治疗经过(如自我保健措施、近期就诊情况、以前接受的其他医疗服务等)     | 1 | 2 | 3 | 4 | 5 |
| 10. 与病人讨论目前疾病对其生活的影响(如生活质量)                     | 1 | 2 | 3 | 4 | 5 |
| 11. 与病人讨论健康的生活方式/疾病预防措施(如疾病危险因素)                | 1 | 2 | 3 | 4 | 5 |
| 12. 避免诱导性提问/命令式提问                               | 1 | 2 | 3 | 4 | 5 |
| 13. 给病人说话的时间和机会(如不轻易打断病人的讲话)/无尴尬停顿              | 1 | 2 | 3 | 4 | 5 |
| 14. 用心倾听(如面朝病人、肯定性的语言、非言语的意见反馈等)                | 1 | 2 | 3 | 4 | 5 |
| 15. 核实/澄清所获得的信息(如复述、询问具体的数量)                    | 1 | 2 | 3 | 4 | 5 |
| <b>信息给予</b>                                     |   |   |   |   |   |
| 16. 解释诊断性操作的理论抑郁(如体格检查、实验室检查等)                  | 1 | 2 | 3 | 4 | 5 |
| 17. 告诉病人他(她)目前身体情况(如体格检查/实验室检查的结果, 解剖学一场/诊断的结果) | 1 | 2 | 3 | 4 | 5 |
| 18. 鼓励病人提问/核实自己的理解/安慰、鼓励病人                      | 1 | 2 | 3 | 4 | 5 |
| 19. 根据病人的理解能力进行适当(语速、音量)调整(如避免使用/解释专业术语)        | 1 | 2 | 3 | 4 | 5 |
| <b>理解病人</b>                                     |   |   |   |   |   |
| 20. 认同病人所付出的努力/所取得的成就/所需要克服的困难(如感谢病人的配合)        | 1 | 2 | 3 | 4 | 5 |
| 21. 体察病人的暗示/配合默契                                | 1 | 2 | 3 | 4 | 5 |
| 22. 表达关心、关注、移情/使病人感到温暖/树立信心                     | 1 | 2 | 3 | 4 | 5 |
| 23. 始终保持尊重的语气                                   | 1 | 2 | 3 | 4 | 5 |
| <b>结束问诊</b>                                     |   |   |   |   |   |
| 24. 问病人是否还有其它的问题需要探讨                            | 1 | 2 | 3 | 4 | 5 |
| 25. 进一步说明下一步的诊治方案                               | 1 | 2 | 3 | 4 | 5 |

© Copyright 1993/2000 Gregory Thomas Makoul- All Rights Reserved.

References

[1] Ha JF, Longnecker N. Doctor-patient communication: a review. *Ochsner J.* 2010;10(1): 38–43.

[2] Brinkman WB, et al. Effect of multisource feedback on resident communication skills and professionalism: a randomized controlled trial. *Arch Pediatr Adolesc Med.* 2007; 161(1):44–9.

[3] Boucher VG, et al. Evaluation of communication skills among physicians: a systematic review of existing assessment tools. *Psychosom Med.* 2020;82(4):440–51.

[4] Makoul G. The SEGUE framework for teaching and assessing communication skills. *Patient Educ Couns.* 2001;45(1):23–34.

[5] Guo A, Wang P. The current state of doctors' communication skills in mainland China from the perspective of doctors' self-evaluation and patients' evaluation: a cross-sectional study. *Patient Educ Couns.* 2021;104(7):1674–80.

[6] Assis-Hassid S, et al. Existing instruments for assessing physician communication skills: are they valid in a computerized setting? *Patient Educ Couns.* 2013;93(3):363–6.

[7] Lundberg KL. What are internal medicine residents missing? A communication needs assessment of outpatient clinical encounters. *Patient Educ Couns.* 2014;96 (3):376–80.

[8] Xiong Y, et al. Development and validation of online textual pediatrician-parent communication instrument based on the SEGUE framework. *Biomed Res Int.* 2019;.2019.

[9] Muddle L, O'Malley CJ, Stupans I. Adaptation and validation of the SEGUE checklist to assess osteopathy students' clinical communication skills. *Int J Osteopath Med.* 2021;42: 100–5.

[10] Wang Y, et al. Doctors' and patients' perceptions of impacts of doctors' communication and empathy skills on doctor–patient relationships during COVID-19. *J Gen Intern Med.* 2022;1–6.

[11] Putnick DL, Bornstein MH. Measurement invariance conventions and reporting: the state of the art and future directions for psychological research. *Dev Rev.* 2016;41: 71–90.

[12] Van de Vijver FJ. Towards a theory of bias and equivalence. 1998;50967(3):41–65.

[13] Vandenberg RJ, Lance CE. A review and synthesis of the measurement invariance literature: suggestions, practices, and recommendations for organizational research. *Organ Res Methods.* 2000;3(1):4–70.

[14] Emerson SD, Guhn M, Gadermann AM. Measurement invariance of the satisfaction with life scale: reviewing three decades of research. *Qual Life Res.* 2017;26:2251–64.

[15] Papayiannis S, et al. Cross-cultural studies. *Encycl Child Behav Dev.* 2011:438–40.

[16] Anakwah N, et al. Cross-cultural differences in eyewitness memory reports. *Appl Cogn Psychol.* 2020;34(2):504–15.

[17] Gong W, et al. The lineage theory of the regional variation of individualism/collectivism in China. *Front Psychol.* 2021;11:596762.

[18] Ling Y, et al. Assessing the measurement invariance of the gratitude questionnaire–5 in Chinese and Italian primary school student samples. *Assessment.* 2012;19(4):506–16.

[19] Wu W, et al. Assessing measurement invariance of the Children's depression inventory in Chinese and Italian primary school student samples. *Assessment.* 2012;19(4):506–16.

[20] Kankaraš M, Moors G. Researching measurement equivalence in cross-cultural studies. *Psihologija.* 2010;43(2):121–36.

[21] Sorrel MA, et al. Cross-cultural measurement invariance in the personality inventory for DSM-5☆. *Psychiatry Res.* 2021;304:114134.

[22] Holland PW, Wainer H. Differential item functioning. *Psychology Press*; 1993.



- [23] Mellenbergh GJ. Item bias and item response theory. *Int J Educ Res.* 1989;13(2): 127–43.
- [24] French BF, Finch WH. Confirmatory factor analytic procedures for the determination of measurement invariance. *Struct Equ Modeling.* 2006;13(3):378–402.
- [25] Wells CS. Assessing measurement invariance for applied research. Cambridge University Press; 2021.
- [26] Bialosiewicz S, Murphy K, Berry T. Do our measures measure up? The critical role of measurement invariance. An introduction to measurement invariance testing: Resource packet for participants, Claremont; 2013.
- [27] Lee ST. Testing for measurement invariance: does your measure mean the same thing for different participants? *APS Obs.* 2018:31.
- [28] Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika.* 1951; 16(3):297–334.
- [29] Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Model Multidiscip J.* 1999;6(1): 1–55.
- [30] Cheung GW, Rensvold RB. Evaluating goodness-of-fit indexes for testing measurement invariance. *Struct Equ Modeling.* 2002;9(2):233–55.
- [31] Pasca L, et al. The equivalence of measures on the connectedness to nature scale: a comparison between ordinal methods of DIF detection. *PloS One.* 2018;13(11):e0207739.
- [32] Wang J, et al. Validation of the Chinese version of the procrastination at work scale. *Front Psychol.* 2021;12:726595.
- [33] Schouten BC, Meeuwesen L. Cultural differences in medical communication: a review of the literature. *Patient Educ Couns.* 2006;64(1–3):21–34.
- [34] Zhang P, et al. Personality inventory for DSM-5 brief form (PID-5-BF) in Chinese students and patients: evaluating the five-factor model and a culturally informed six-factor model. *BMC Psychiatry.* 2021;21(1):1–12.
- [35] Van Doren N, Zainal NH, Newman MG. Cross-cultural and gender invariance of emotion regulation in the United States and India. *J Affect Disord.* 2021;295:1360–70.
- [36] Atienza González FL, et al. Factorial invariance of the satisfaction with life scale in adolescents from Spain and Portugal. *Psicothema.* 2016;28(3):353–8.
- [37] Arrindell WA, et al. Measurement invariance and construct validity of the satisfaction with life scale (SWLS) in community volunteers in Vietnam. *Int J Environ Res Public Health.* 2022;19(6):3460.
- [38] Zager Kocjan G, et al. Measurement invariance of the subjective happiness scale across countries, gender, age, and time. *Assessment.* 2022;29(4):826–41.
- [39] Zhou Y, et al. Cross-cultural measurement invariance of scales assessing stigma and attitude to seeking professional psychological help. *Front Psychol.* 2019;10:1249.
- [40] Chan MH-M, Gerhardt M, Feng X. Measurement invariance across age groups and over 20 years' time of the negative and positive affect scale (NAPAS). *Eur J Psychol Assess.* 2019;36:537–44.