

Development and Cross-National Validation of a French Version of the Jefferson Scale of Empathy for Students

Margaret W. Gerbase¹ , Marie-Paule Gustin², Nadia Bajwa³, Milena Abbiati¹, and Anne Baroffio¹

Evaluation & the Health Professions
2022, Vol. 45(3) 288–296
© The Author(s) 2021



Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/01632787211033330
journals.sagepub.com/home/ehp



Abstract

Empathy is a multifaceted personal ability combining emotional and cognitive features modulated by cultural specificities. It is widely recognized as a key clinical competence that should be valued during professional training. The Jefferson Scale of Empathy for medical students (JSE-S) has been developed for this purpose and validated in several languages, but not in French. The aims of this study were to gather validity evidence for a newly developed version of the JSE-S and compare it between two French-speaking contexts. In total, 1,433 undergraduate medical students from the universities of Lyon (UL), France and Geneva (UG), Switzerland participated in the study completing the JSE-S in French. Total and partial scores of the three subscales (“perspective taking,” “compassionate care” and “walking in patient’s shoes”) were calculated for each site. Construct validity of the JSE-S was analyzed considering three sources of evidence: content, internal structure and relations to other variables. A first-order Confirmatory Factor Analysis using structural equation modeling examined the three latent variables of the JSE-S subscales. Cronbach’s α coefficients were 0.75 (UG) and 0.81 (UL). The items’ discrimination power ranged between 0.29 and 1.60 (median effect size of 1.24). The overall correlations between items and total or partial scores derived from the latent JSE-S subscales were consistently similar in both study sites. Findings of this study confirm the latent structure of the JSE-S in French and its cross-national reproducibility. The comparable underlying structure of the questionnaire tested in two distinct French-speaking contexts endorses the generalizability of its measure.

Keywords

empathy, validation, psychometric properties, cross-cultural

Empathy is widely recognized as a key competence for clinical practice. A consistent body of evidence has shown that an empathic disposition positively influences communication between healthcare professionals and patients, leading to better compliance, and having a substantial impact on clinical outcomes (Quince et al., 2016). Evidence also indicates that empathy is a multifaceted personal attribute that involves emotional, cognitive and motivational features, which are modulated by social and cultural specificities (Hojat et al., 2001, 2002, 2003). For the purpose of the present study we adopted the conceptual definition of empathy applied to the context of patients’ care and health professionals’ training described by Hojat et al. (2002), as a multidimensional construct encapsulating predominantly a cognitive dimension to understand patient’s experiences and concerns combined with a capacity to communicate this understanding and an intention to help.

In spite of its conceptual complexity involving innate and acquired abilities, it is generally agreed upon that empathy has a considerable positive impact on patients’ clinical outcomes and should therefore be strengthened during clinical training under focused supervision (Benbassat & Bauml, 2004; Hojat et al., 2013). Recently, there has been a growing commitment

of health educators toward the development of empathy among medical trainees, knowing that an empathic behavior will impact on the doctor-patient relationship and professionalism of students’ future practice. Along this line, empathy is currently considered a valued clinical competency for appraisal in medical school admission processes and during medical training (Hojat, 2014; Hojat et al., 2002). However, while several instruments have been developed for the appraisal of empathy among healthcare professionals, few have been designed for the assessment of empathy of medical students during their clinical training. One of these is the Jefferson Scale of Empathy (JSE), a well-known instrument that has been firstly developed

¹Faculty of Medicine, University of Geneva, Switzerland

²INSERM, Paris, France

³Department of Women-Children-Teenagers, University Hospitals of Geneva, Switzerland

Corresponding Author:

Margaret W. Gerbase, Faculty of Medicine, University of Geneva, 1, rue Michel Servet, Geneva 1211, Switzerland.
Email: margaret.gerbase@unige.ch

to assess empathy in healthcare professionals (JSE-HP) and physicians (JSPE), whose authors further developed a version for the assessment of empathy in medical students (JSE-S) (Hojat et al., 2002; Hojat & LaNoue, 2014). All JSE questionnaires were validated in English and its internal latent structure was confirmed underlining three main dimensions: “perspective taking,” “compassionate care” and “walking in patient’s shoes” (Hojat et al., 2002; Hojat & LaNoue, 2014; Tavakol et al., 2011).

Applying standardized questionnaires to different educational, professional, cultural or language contexts, requires confirmation of the instruments’ internal validity. The JSE questionnaires have been largely validated in several languages, notably the version for healthcare professionals (Kataoka et al., 2009; Paro et al., 2012; Preusche & Wagner-Menghin, 2013; Spasenoska et al., 2016; Vallabh, 2011; Williams et al., 2013; Zenasni et al., 2012). Results of these studies confirmed the latent structure previously reported for the English version of the questionnaire. However, a validated version for French-speaking medical students has not been made available so far. In addition, it has been claimed that there is insufficient evidence of the cross-cultural aspects in language validation studies of the JSE (Williams & Beovich, 2019).

A number of studies have examined the relationship of empathy with patterns of social emotions, beliefs and behaviors. Prosocial attitudes and other-oriented response are part of the empathy construct, which can evolve over time. Language is one of the foundation elements that unifies a society and strengthens its values, but it is not the only one. Cultural factors such as empathic parenting, social identity, individual values and educational systems have been associated with the development of empathy (Silke et al., 2018; Szanto & Kruger, 2019; Taylor et al., 2013). On the other hand, empathy embodies a social ideal concept that goes far beyond language and encompasses a collective dynamic that characterizes a geographic region’s population and may or not transcend territorial boundaries. Empathy studies comparing groups pertaining to collectivistic communities with those from more individualistic societies revealed inconclusive (Chopik et al., 2017). Moreover, contrasted empathy scores were found comparing groups sharing the same native language but living in different countries (Alcorta-Garza et al., 2016). Remarkably, in the latter study the lowest JSE scores were observed in the group of participants who had left their country of birth to another of same language. While reasons for these findings are probably multiple, we can speculate whether cultural and motivational aspects beyond language are involved.

To ensure full validity of an instrument to measure a socio-culturally mediated construct such as empathy requires not only cross-language replications but also cross-cultural confirmation. At one side, translation implies challenges to adequately reflect the content and meanings of the instrument-source into the targeted one. On the other hand, cultural nuances across countries or even regions sharing the same language need to be considered. In this study we

integrated the theoretical framework described above in a collaboration between two medical schools from different but geographically close countries in a region sharing the same language. The lack of a validated JSE-S questionnaire in French was a motivational propulsor for researchers from both universities to pursue this work. The aims of this study were to gather construct validity evidence for a newly developed French version of the JSE-S questionnaire and to examine its cross-national validity using Confirmatory Factor Analysis with medical students from two French-speaking countries.

Method

Study Design and Settings

This study was designed as a cross-sectional survey including medical students from two French-speaking universities in different countries. Students from the University of Geneva (UG) in Switzerland and the University of Lyon (UL) in France were invited to participate in the study, which required self-completion of a series of standardized questionnaires at their own site. Previous work by our group detailed the preclinical curricula and the educational formats of the two universities collaborating in this study (Gustin et al., 2018). Briefly, both medical schools have a similar six-year curriculum divided into preclinical and clinical years (1st–3rd and 4th–6th, respectively) and a final licensing exam. However, curricula and instructional formats differ between the two institutions. Geneva has an integrated preclinical curriculum built on Harden’s ladder framework (2000), composed of a variety of teaching methodologies (i.e., lectures, problem-based learning, small groups discussions, simulated clinical activities and early exposure to the clinical environment). Lyon is characterized by a more traditional curriculum, which is based on lectures and tutorials for large classes of students. The clinical curricula of both sites are based on rotations in clinical clerkships.

The study was approved by the Ethical Committee on Human Research of the University of Geneva, Switzerland and the Rectorate of the University of Lyon, France. Agreements for students’ participation were requested and informed consents were given by all participants.

Participants and Data Collection

Out of 1,627 students invited for the study, 1,433 agreed to participate and completed the empathy questionnaires. Excluded from the analyses were 11 students: nine students who did not complete more than four items of the JSE-S ($n = 4$ in UG and 5 in UL) and two students (UG) who did not indicate their gender. A total of 1,433 students’ responses to the questionnaires were analyzed ($n = 739$, UG; $n = 694$, UL).

Data used in this study were derived from answers to the JSE-S questionnaire applied to two cohorts of students in years 1, 4 and 5 in UG and two independent cohorts of students in years 2 and 4 in UL. The collection of data started in 2011 when the first cohort enrolled the 6-year curriculum of medical

studies and lasted until 2017 when the last cohort finished medical studies. Questionnaires were completed either on paper (UG) or online (UL). The overall rates of participation in the study were 87% and 90% in UG and UL, respectively.

Empathy Definition and Instruments

Students participating in the study answered questions on demographic data and the JSE-S questionnaire in French. The JSE-S is a 20-item standardized self-reporting questionnaire, consisting of a similar version of the original Jefferson Scale of Physician's Empathy where wording was adapted for the assessment of students' perception of the importance of empathy in the medical profession (Kataoka et al., 2009). The response to each of the items is based on a 7-point Likert scale (fully disagree = 1 to fully agree = 7). The questionnaire includes 10 reverse-scored items and total scores range from 20 (lowest) to 140 (highest). Partial scores were calculated by the sum of the items scores contained in each of the three JSE-S subscales: "perspective taking," "compassionate care" and "walking in patient's shoes."

To examine the relationship of the JSE-S to an external variable (convergent validity), the Empathy Quotient (EQ) was administered to a sample of 852 students from both sites agreeing to complete this additional questionnaire at the same time as the JSE-S. Several empathy instruments are available that could be used for the purpose of external validity in our study. The choice of the EQ was mainly based on the fact that this questionnaire was conceived for the assessment of empathy in young adults (Lawrence et al., 2004) similar in age to our students and has been validated in French (Berthoz et al., 2008). In addition, the EQ questionnaire has a shorter version containing 40 questions, which was suitable for our study where students were asked to complete a series of other questionnaires regarding personal characteristics and perception of the educational environment. Finally, the items of the EQ were elaborated on a direct first-person approach, whereas the JSE' items were formulated on a more indirect mode representing what students think "should" be the behavior of a doctor. The students' rate of participation filling the second questionnaire of empathy was 60%.

Procedure for the Translation and Adaptation of the JSE-S Version to the French Language

Initially, the original JSE-S questionnaire in English was translated into French by the team; then, an external reviewer checked for wording and grammar accuracy. The questionnaire was then back translated by a native English speaker and further back translated into French by a native speaker. Finally, the research team performed back translations to ensure the content and conceptual equivalence of the translation (Helmlich et al., 2017). Minimal wording corrections warranted the suitability of the questionnaire for the two study sites, and the final version was approved by all parties.

Analyses

Demographic data included age, gender, site of medical school and year of study. Descriptive analyses were performed for demographic data and the JSE-S total scores. Results are reported as means, standard deviations, median and range for the entire sample of students and after stratification by site and gender. A two-way ANOVA was performed with site and gender as factors. Their interaction effect on the JSE-S total score was analyzed.

Construct validity of the JSE-S was analyzed considering the widely accepted unified validity framework operationalized by Downing (2003). For the purpose of this study, we considered three consequential sources of validity evidence: content, internal structure and relations to other variables.

Content. The content was assessed by the analysis of the homogeneity of items' distribution in the three JSE-S subscales, as described in previous work by the questionnaire developers (Hojat & Gonnella, 2015; Kataoka et al., 2009).

Internal structure. The internal structure of the JSE-S in French was examined by:

1. The calculation of the JSE-S item-total score correlations and the correlations between the JSE-S subscales. Pearson's correlation coefficients were calculated for each item by using its score and the total scale score, after omitting that specific item from the total score;
2. The estimation of the discrimination power of each questionnaire item. For this purpose, we adopted the procedure described by Hojat and colleagues (Hojat et al., 2002; Kataoka et al., 2009). As such, the total dataset was divided in three groups and two groups were retained: the first corresponded to the bottom tertile (JSE-S score < 106, n = 441) and the second to the top tertile (JSE-S score > 116, n = 456). Then the discrimination power was estimated by the difference between the scores in the top-tertile and the bottom-tertile groups divided by the pooled standard deviation of the item scores in both groups (Cohen, 1988). Accordingly, we considered that the effect size < 0.30 was small, between 0.30 and 0.70 moderate, and > 0.70 large;
3. The reliability, which was measured by: a) the calculation of the Cronbach's α coefficients for the overall JSE-S scores and for the three subscales' scores of the questionnaire, before and after the stratification of data by site; and b) the reproducibility of the JSE-S scores verified by test-retest performed on a subset of 55 students who completed the same questionnaires within a 3-month interval. Data from the retest were not included in the overall analyses;
4. The factorial structure of the three underlying domains of the JSE-S in French was studied by applying a first-order confirmatory factor analysis (CFA) using

structural equation modeling (SEM). For this analysis, we considered the three latent variables corresponding to the three JSE-S subscales: “perspective taking” (items 2, 4, 5, 9, 10, 13, 15, 16, 17 and 20), “compassionate care” (items 1, 7, 8, 11, 12, 14, 18 and 19) and “walking in patient’s shoes” (items 3 and 6). Geomin rotation and the Mplus option MLR for maximum likelihood estimation with robust standard errors were used. The MLR option is recommended in case of non-normal distribution of item scores to handle missing values. The root mean square error for approximation (RMSEA) was used to confirm the latent variable structure of the French JSE-S questionnaire. A RMSEA less than 0.06 indicates good model fit (Hu & Bentler, 1999). To compare our results with those of LISREL SEM software users, two of the most commonly used incremental of fit were reported as well: the normed comparative fit index (CFI) and the non-normed Tukey-Lewis index (TLI), which were adjusted for model complexity. These fit indices measure the proportionate improvement in fit of a hypothesized model compared with a null model having no underlying component. They are similar to the AGFI and GFI indices of LISREL respectively.

Relations to other variables. The relationship to an external variable (i.e., the correlation between the JSE-S and the EQ scores) was used to examine convergent validity evidence. It was tested in the overall sample of students and in the two groups according to site (UG and UL). Using multiple linear models, we tested how EQ and covariables predicted the JSE-S total score to determine the covariables that acted on JSE-S with significant interaction with EQ. These covariables with significant interaction were kept in the groups’ definition.

Data were analyzed using the Mplus version 7.11 (available at <https://www.statmodel.com/>) for structural equation modeling (SEM) and using R version 3.3.1 (available at <https://cran.r-project.org/>) for other statistical analyses. Missing data (0.8%) were handled by Mplus in case of SEM. A p -value ≤ 0.05 was considered statistically significant.

Results

Descriptive Analyses

Age and sex distribution differed between UG and UL. Taking the data from year 4, the one when students from both sites completed the questionnaires, ages were 23.6 ± 1.7 and 22.6 ± 1.1 ($p < 0.001$), respectively for UG and UL. Compared to men, the proportion of women was higher in both sites (64.4% at UG and 57.5% at UL ($p = 0.007$), though lower in UL. Table 1 shows the descriptive results for the entire sample of students and after stratification by gender and site. Overall differences in JSE-S scores between sites and comparing men and women were statistically significant ($p < 0.0001$). Statistical

Table 1. JSE-S Scores Calculated for the Entire Sample of Students and After Stratification by Gender and Study Sites.

| Sample | Groups | N | Mean (SD) | Median | Range |
|--------|--------|-------|--------------|--------|--------|
| All | All | 1,433 | 110.0 (11.9) | 111 | 35–140 |
| | UG | 739 | 112.9 (10.0) | 114 | 76–140 |
| | UL | 694 | 106.9 (13.0) | 108 | 35–134 |
| | Women | 875 | 112.3 (10.5) | 113 | 47–140 |
| | Men | 558 | 106.5 (13.2) | 108 | 35–132 |
| UG | Women | 476 | 114.4 (9.3) | 115 | 76–140 |
| | Men | 263 | 110.1 (10.7) | 111 | 77–132 |
| UL | Women | 399 | 109.7 (11.2) | 111 | 47–134 |
| | Men | 295 | 103.2 (14.3) | 105 | 35–131 |

Note. JSE-S: Jefferson Scale of Empathy for medical students; UG: University of Geneva; UL: University of Lyon. All pairwise comparisons were statistically significant ($p < 0.0001$).

Table 2. Correlations Between Items and Total Scores ($n = 1,433$), and Effect Size Estimates of Item Discrimination Indices for the 20 Items Loaded in the Three Domains of the JSE-S ($n = 897$).

| Underlying Components | Items | Item-Total Scores Correlations | Discrimination Effect Size Index |
|----------------------------|-------|--------------------------------|----------------------------------|
| Perspective taking | 2 | 0.50 | 1.3 |
| | 4 | 0.43 | 1.1 |
| | 5 | 0.16 | 0.59 |
| | 9 | 0.41 | 1.4 |
| | 10 | 0.53 | 1.5 |
| | 13 | 0.56 | 1.6 |
| | 15 | 0.49 | 1.6 |
| | 16 | 0.59 | 1.5 |
| | 17 | 0.33 | 1.2 |
| | 20 | 0.60 | 1.6 |
| Compassionate care | 1 | 0.36 | 1.2 |
| | 7 | 0.49 | 1.3 |
| | 8 | 0.53 | 1.4 |
| | 11 | 0.49 | 1.2 |
| | 12 | 0.51 | 1.2 |
| | 14 | 0.50 | 1.3 |
| | 18 | −0.0061 | 0.29 |
| 19 | 0.22 | 0.64 | |
| Walking in patient’s shoes | 3 | 0.091 | 0.41 |
| | 6 | 0.048 | 0.47 |

Note. JSE-S: Jefferson Scale of Empathy for Medical Students.

differences were also observed when comparing the JSE-S scores between sites in the group of men ($p < 0.0001$) and in women ($p < 0.0001$). In general, the range of scores was larger in UL compared to UG.

JSE-S Item-Total Scores Correlations

The item-total score correlations ranged from -0.0061 for item 18, which also had the lowest mean, to 0.60 for item 20 (Table 2). All correlations were positive and significant ($p < 0.01$), except for item 18 ($r = -0.0061$, $p = 0.82$) and item 6 ($r = 0.048$, $p = 0.07$). The median correlation between items and

Table 3. Cronbach's α Coefficients Calculated for the JSE-S Total Score and the Three Subscale Scores for the Total Sample of Students and After Stratification by Site.

| | Groups | n | # Items | Cronbach's α | 2.5% CI | 97.5% CI |
|----------------------------|--------|-------|---------|---------------------|---------|----------|
| Total JSE-S score | ALL | 1,433 | 20 | 0.80 | 0.77 | 0.82 |
| | UG | 739 | 20 | 0.75 | 0.72 | 0.78 |
| | UL | 694 | 20 | 0.81 | 0.78 | 0.84 |
| Perspective taking | ALL | 1,433 | 10 | 0.79 | 0.76 | 0.80 |
| | UG | 739 | 10 | 0.73 | 0.70 | 0.76 |
| | UL | 694 | 10 | 0.79 | 0.76 | 0.82 |
| Compassionate care | ALL | 1,433 | 8 | 0.67 | 0.64 | 0.71 |
| | UG | 739 | 8 | 0.62 | 0.56 | 0.66 |
| | UL | 694 | 8 | 0.70 | 0.66 | 0.74 |
| Walking in patient's shoes | ALL | 1,433 | 2 | 0.58 | 0.52 | 0.63 |
| | UG | 739 | 2 | 0.56 | 0.48 | 0.63 |
| | UL | 694 | 2 | 0.60 | 0.52 | 0.66 |

Note. JSE-S: Jefferson Scale of Empathy for Medical Students; UG: University of Geneva (Switzerland); UL: University of Lyon (France); n: sample size; #items: number of items; Cronbach's α : coefficients of reliability; CI: confidence intervals (95% CI: 2.5% = lower limits, 97.5% = upper limits).

total score was 0.49. The non-left skewed items were those with the lowest correlations.

Discrimination Power of the JSE-S Items

Table 2 shows the discrimination effect sizes, which ranged between 0.29 (item 18, lowest mean and lowest item-total score correlation) and 1.6 (item 20, highest item-total score correlation). Items 3, 6 and 18 showed small effect sizes or in the lower limit of moderate, which was consistent with their low item-total score correlation. The median effect size was 1.24 and the first quartile of the effect size was 1.00 (> 0.70), thus showing that 75% of the items had an important discrimination effect.

Reliability and Reproducibility of the JSE-S

Table 3 depicts the Cronbach's α coefficients of reliability for the JSE-S total scores. α coefficients were 0.75 (UG) and 0.81 (UL) with lower 95% CI (confidence interval) limits found higher than 0.70 in all groups. Stratified by site and subscales, the Cronbach's α coefficients ranged from 0.73 (UG) to 0.79 (UL), from 0.62 (UG) to 0.70 (UL) and from 0.56 (UG) to 0.60 (UL) for "perspective taking," "compassionate care" and "walking in patient's shoes," respectively. As anticipated, the α coefficients decreased with the number of items, but they remained acceptable notably for the first two subscales.

Test-retest of the total JSE-S scores showed a weak and non-significant difference of 2.5 points over an interval of 120 points of the total JSE-S score, corresponding to a 2.1% difference between the questionnaires applied on two occasions ($p = 0.07$).

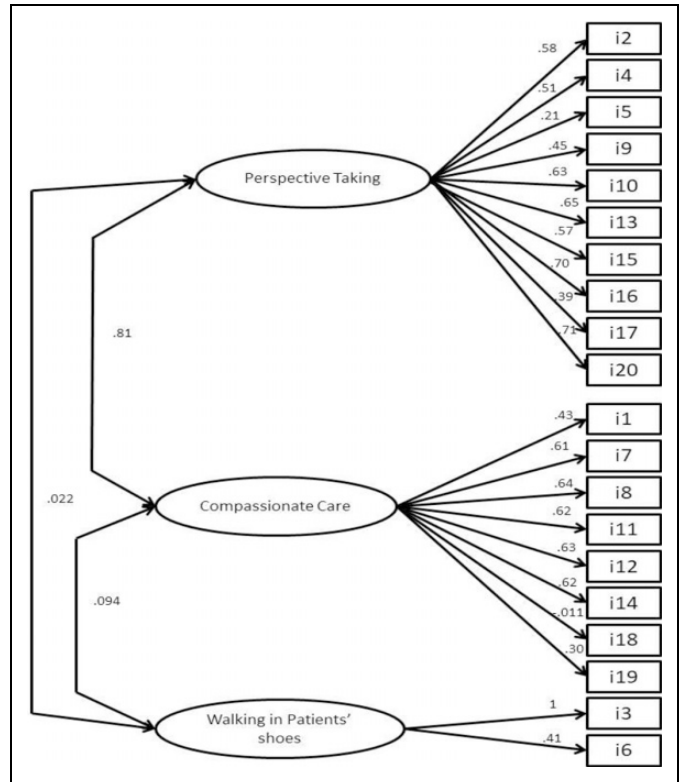


Figure 1. Overall results of first-order confirmatory factor analysis model of the JSE-S with three latent variables (n = 1,433). Note. JSE-S: Jefferson Scale of Empathy for Medical Students. The numbers correspond to the standardized coefficients of the model. The latent variables and the items are represented by an ellipse and a square, respectively. The single-directed arrows linking latent variables to items give the correlation between that item and the latent variable. The double-directed arrows between two latent variables give their correlation.

Factorial Validity

The results of the CFA showed that the first order three-factor model provided an acceptable fit (RMSEA = 0.050 < 0.06). The CFI and the TLI were 0.876 and 0.860, respectively. In the whole study sample, the residual variance of item 3 was initially negative and not significantly different from zero; this residual variance was then fixed to zero.

The correlations between the items and their latent variable ranged from 0.21 (item 5) to 0.71 (item 20) for "perspective taking," from -0.011 (item 18) to 0.64 (item 8) for "compassionate care" and from 0.41 (item 6) to 1.0 (item 3) for "walking in patient's shoes" (Figure 1). The first two latent variables were highly correlated (0.81), but the third variable "walking in patient's shoes" was not or weakly correlated with the other two, i.e. 0.022 ($p = 0.49$) with "perspective taking" and 0.094 ($p = 0.003$) with "compassionate care." Item 18 was negative and not significantly correlated with "compassionate care" (i.e., -0.011; $p = 0.75$).

The three-factor model obtained after suppression of item 18 reported a similar fit (RMSEA = 0.052, CFI = 879, TLI = 0.862). The two-factor model obtained after suppression of the

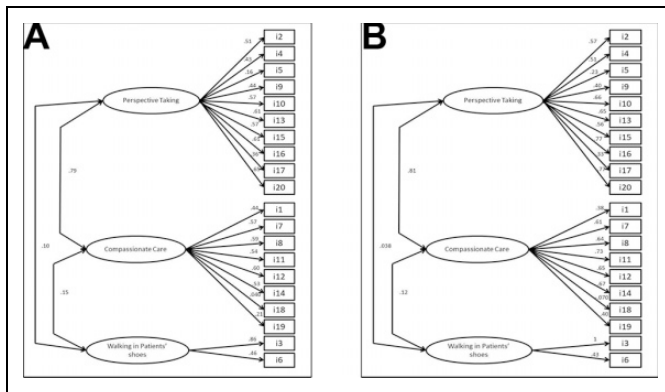


Figure 2. First-order confirmatory analysis model for the JSE-S with three latent variables in the two study sites. *Note.* Results are presented for medical students from the University of Geneva (A; $n = 739$) and medical students from the University of Lyon (B; $n = 694$). JSE-S: Jefferson Scale of Empathy for Medical Students. The numbers correspond to the standardized coefficients of the model. The latent variables and the items are represented by an ellipse and a square, respectively. The single-directed arrows linking latent variables to items give the correlation between that item and the latent variable. The double-directed arrows between two latent variables give their correlation.

third latent variable “walking in patient’s shoes” exhibited an acceptable fit as well (RMSEA = 0.053, CFI = 0.878, TLI = 0.861). As the two latter models displayed similar fit, we concluded that the three-factor model with 20 items could be retained.

Figure 2 shows the three-factor models in each French-speaking site, namely UG and UL. Both models had acceptable fits with a RMSEA of 0.042 and 0.056, respectively. The residual variance of item 3 was fixed to zero only in UL. The correlations between the latent variables were equivalent in the two sites, being high between “perspective taking” and “compassionate care,” and low between “walking in patient’s shoes” and the two others. Correlations between items and partial subscale scores were found to be similar at the two sites. Item 18 was not significantly correlated with “compassionate care” at both sites.

Relationship Between the JSE-S and the External Variable EQ scores

The comparison of the JSE-S scores between students who completed or not the EQ questionnaire was not statistically relevant (109.6 ± 12.8 and 110.6 ± 10.5 , respectively; $p = 0.11$). Pearson’s correlation between the total JSE-S scores and the EQ scores was 0.45 (95% IC 0.39, 0.50; $p < 0.0001$). Furthermore, correlations between the JSE-S subdomains and the EQ were 0.44 (95% IC 0.39, 0.50; $p < 0.0001$), 0.34 (95% IC 0.27, 0.39; $p < 0.0001$) and 0.09 (95% IC 0.024, 0.16; $p < 0.008$), respectively for “compassionate care,” “perspective taking” and “walking in patient’s shoes.”

After stratifying by site, Pearson’s correlations between the total JSE-S scores and the EQ scores were 0.35 (95% IC 0.27,

0.43) and 0.44 (95% IC 0.36, 0.52), respectively for UG and UL.

Discussion

This study was conceived to fill a gap in the literature by providing a validated version of the JSE-S which could be used by native French speakers around the world. In addition, we aimed at examining the generalizability of the underlying latent structure of the JSE-S version in French, by applying the new questionnaire to students from two French-speaking academic environments in different countries. The analyses showed consistent results when compared to the original English version (Hojat et al., 2018; Kataoka et al., 2009). Moreover, a similar factorial structure was found between the two French-speaking sites. To our knowledge, this is the first cross-national confirmatory factor analysis of the JSE-S underlying constructs comparing two samples of undergraduate medical students in the French language across countries.

Extensive literature is available on the validation of the JSPE questionnaire, and confirmatory factorial analyses endorsed its underlying internal structure and the stability of the three latent variables originally described in the English version (Kataoka et al., 2009; Paro et al., 2012). In addition, several studies validated the JSE-S in different languages, confirming its factorial structure, except in French (Spasenoska et al., 2016; Williams et al., 2013; Zenasni et al., 2012). In line with these reports, we were able to corroborate its generalizability and expand findings by further showing a consistent replication of the internal structure of the JSE-S in two French-speaking countries.

Reports comparing the validation of the JSE across countries with the same language are scarce. Alcorta-Garza and colleagues (2016) examined the psychometric features of the JSPE for physicians comparing data from Spanish-speaking professionals in Spain and countries in Latin America. Their results supported the factorial structure found in the original English version of the JSPE but highlighted significant differences across countries in the magnitude of scores. By comparison, we found similar results with contrasted JSE-S scores between sites, but confirmed the factorial structure of the questionnaire at the two French-speaking sites. Overall, the JSE-S scores found in our study were similar to previous reports, and the observed gender differences are in line with the literature showing that women scored higher, as compared to men at the two study sites (Casas et al., 2017; Ferreira-Valente et al., 2017; Kataoka et al., 2009; Quince et al., 2016).

The results showed a relatively higher reliability in one site (UL), as compared to the other. While we cannot ensure that the latter observed differences were related to students’ characteristics in the two countries, we speculate whether better reliability could partly be derived from the online format applied by one site (UL), which had a slightly higher rate of participation over the other site.

Another aspect of the results that caught our attention was the relatively modest correlation between the JSE-S and the EQ

questionnaires, a finding that has also been pointed out by others. This could be linked to the overall construct and the dimensions of empathy assessed by the two instruments. While the JSE-S considers more the cognitive domain (Smith et al., 2017), the EQ addresses more the affective component of empathy. In addition, the items of the EQ were elaborated to assess the person's own feelings regarding empathy, whereas the JSE' items indicate what a person believes is an empathic behavior. These contrasted features could explain part of our findings, eventually indicating that a single-instrument approach to measure empathy may not be sufficient. In a study analyzing the dimensional factors of the JSE-S focusing on medical students from five countries with different languages, Costa and co-authors (2017) corroborated the cross-cultural validity and stability of the scale, but they highlighted the weak comparability between the JSE-S and the IRI (Interpersonal Reactivity Index) questionnaires. Reasons for this finding have been imputed to the complexity of the empathy definition, the challenges of assessing it in different language contexts, the disproportion between cognitive and affective domains contained in different questionnaires and the recently evoked geo-sociocultural characteristics including language, which might impact the way individuals value empathy and how they respond to specific questionnaire items (Helmlich et al., 2017; Ponnampereuma et al., 2019).

One can be puzzled by the myriad of components that might influence the development of a person's empathic disposition. Available evidence shows that such factors shaping individuals' emotions may vary widely from parental support during childhood, personality, coping and to newly identified brain regions linked to emotions (Chopik et al., 2017; Decety, 2020). Moreover, it has been upheld that culture gives structure to how people build experience and feel and express compassion, which are central elements of empathy (Koopman-Holm & Tsai, 2017). Along this line, studies comparing western and eastern populations have shown cross-cultural differences in the way people feel and react to others' emotions, thus highlighting the geo-cultural aspect potentially linked to empathy (Andersen et al., 2020; Hollan, 2012). Our study was not designed to examine the specific cultural aspects of empathy, and we can only conjecture whether, beyond language, some of our converging findings might indicate cultural similarities between students across the two countries.

Strengths of this study are the systematic and standardized collection of data from a large sample of medical students derived from distinct cultural environments in two French-speaking countries. However, the study also has some limitations. First, students' academic and individual characteristics, which might have a potential impact on empathy (e.g., workload, curricula, stress) were not in the scope of this study. Second, questionnaires were completed on a self-reported basis adopting different formats in each setting (paper vs. online). Therefore, we cannot ensure that all students completed the task with equal diligence. However, the low rates of missing values indicate that students were generally well disposed to fill in the questionnaires irrespectively of site. Along this line,

the difference in classes that took the survey at each location could have played a role in the contrasted JSE-S scores observed between the two sites; however, the higher empathy scores observed in UG probably reflected the higher rates of women with significantly higher JSE-S scores at this site, compared to UL. Third, our data derived from students enrolled on a longitudinal basis and analyses did not consider their level of advancement in the curriculum. Available evidence had reported relevant differences in the JSE scores and factorial analysis between preclinical and clinical years, but studies have been inconsistent (Hojat et al., 2009; Quince et al., 2011; Roff, 2015; Stanfield et al., 2016; Szanto & Kruger, 2019). In a previous work by our group, we showed that the JSE-S scores remained longitudinally stable in the majority of our students (Piumatti et al., 2020). Finally, the discrimination effect size of certain items of the JSE-S were below a minimum level, potentially affecting the results. In this sense, notably item 18 of the subdomain "compassionate care" displayed a very low effect size. Previous work had demonstrated the low discrimination effect of this item (Alcorta-Garza et al., 2016; Kataoka et al., 2009). We therefore performed a sensitivity analysis after suppression of the item with no substantial effect on the model fit.

In conclusion, findings of this study confirm the underlying structure of the JSE-S questionnaire in French, endorsing previous reports in other language contexts. Moreover, the similarities observed in the factorial structure of the JSE-S applied to two distinct French-speaking academic contexts highlights its generalizability. Beyond the language, future research might expand the diverse geo-sociocultural contexts and further discuss definitions and the methodological adequacy of the approaches so far employed to assess the construct of empathy (Hall & Schwartz, 2019). On the other hand, the relative low comparability among instruments meant to measure the same construct may indicate the potential complementarity of different approaches needed to assess a concept as complex as empathy.


Declaration of Conflicting Interests

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

Funding

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

ORCID iD

Margaret W. Gerbase  <https://orcid.org/0000-0001-5185-7811>

References

- Alcorta-Garza, A., San-Martin, M., Delgado-Bolton, R., Soler-González, J., Roig, H., & Vivanco, L. (2016). Cross-validation of the Spanish HP-version of the Jefferson Scale of Empathy confirmed with some cross-cultural differences. *Frontiers in Psychology, 7*, 1–9.
- Andersen, F. A., Johansen, A. S. B., Sondergaard, J., Andersen, C. M., & Hvidt, E. A. (2020). Revisiting the trajectory of medical

- students' empathy, and impact of gender, specialty preferences and nationality: A systematic review. *BMC Medical Education*, <https://doi.org/10.1186/s12909-020-1964-5>
- Benbassat, J., & Baumal, R. (2004). What is empathy and how can it be promoted during clinical clerkships. *Academic Medicine*, *79*(9), 832–839.
- Berthoz, S., Wessa, M., Kedia, G., Wicker, B., & Grèzes, J. (2008). Cross-cultural validation of the empathy quotient in a French-speaking sample. *Canadian Journal of Psychiatry*, *53*(6), 37–45.
- Casas, R. S., Xuan, Z., Jackson, A. H., Stanfield, L. E., Harvey, N. C., & Chen, D. C. (2017). Associations of medical student empathy with clinical competence. *Patient Education and Counselling*, *100*, 742–747.
- Chopik, W. J., O'Brien, E., & Konrath, S. H. (2017). Differences in empathic concern and perspective taking across 63 countries. *J Cross-Cultural Psychology*, *48*(1), 23–28.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Academic Press.
- Costa, P., Carvalho-Filho, M. A., Schweller, M., Thiemann, P., Salgueira, A., Benson, J., Costa, M. J., & Quince, T. (2017). Measuring medical students' empathy: Exploring the underlying constructs of and associations between two widely used self-report instruments in five countries. *Academic Medicine*, *92*(6), 860–867.
- Decety, J. (2020). L'empathie en médecine. *Annales Médico-Psychologiques*, *178*, 197–206.
- Downing, S. M. (2003). Validity: On the meaningful interpretation of assessment data. *Medical Education*, *37*, 830–837.
- Ferreira-Valente, A., Monteiro, J. S., Barbosa, R. M., Salgueira, A., Costa, P., & Costa, M. J. (2017). Clarifying changes in student empathy throughout medical school: A scoping review. *Advances in Health Sciences Education*, *22*(5), 1293–1313.
- Gustin, M. P., Abbiati, M., Bonvin, R., Gerbase, M. W., & Baroffio, A. (2018). Integrated problem-based learning versus lectures: A path analysis modelling of the relationships between educational context and learning approaches. *Medical Education Online*, *23*, 1–12.
- Hall, J. A., & Schwartz, R. (2019). Empathy present and future. *Journal of Social Psychology*, *159*(3), 225–243.
- Harden, R. M. (2000). The integration ladder: A tool for curriculum planning and evaluation. *Medical Education*, *34*(7), 551–557.
- Helmich, E., Cristancho, S., Diachun, L., & Lingard, L. (2017). 'How would you call this in English?' Being reflective about translations in international, cross-cultural qualitative research. *Perspectives in Medical Education*, *6*, 127–132.
- Hojat, M. (2014). Assessments of empathy in medical school admissions: What additional evidence is needed. *International Journal of Medical Education*, *5*, 7–10.
- Hojat, M., Axelrod, D., Spandorfer, J., & Mangione, S. (2013). Enhancing and sustaining empathy in medical students. *Medical Teacher*, *35*, 996–1001.
- Hojat, M., DeSantis, J., Shannon, S. C., Mortensen, L. H., Speicher, M. R., Bragan, L., LaNoue, M., & Calabrese, L. H. (2018). The Jefferson Scale of Empathy: A nationwide study of measurement properties, underlying components, latent variable structure, and national norms in medical students. *Advances in Health Sciences Education*, *23*, 899–920.
- Hojat, M., & Gonnella, J. S. (2015). Eleven years of data on the Jefferson scale of empathy-medical student version (JSE-s): Proxy norm data and tentative cutoff scores. *Medical Principles and Practice*, *24*, 344–350.
- Hojat, M., Gonnella, J. S., Mangione, S., Nasca, T. J., & Magee, M. (2003). Physician empathy in medical education and practice: Experience with the Jefferson scale of physician empathy. *Seminars in Integrative Medicine*, *1*(1), 25–41.
- Hojat, M., Gonnella, J. S., Mangione, S., Nasca, T. J., Veloski, J. J., Erdmann, J. B., Callahan, C. A., & Magee, M. (2002). Empathy in medical students as related to academic performance, clinical competence and gender. *Medical Education*, *36*, 522–527.
- Hojat, M., Gonnella, J. S., Nasca, T. J., Mangione, S., Vergare, M., & Magee, M. (2002). Physician empathy: Definition, components, measurement, and relationship to gender and specialty. *American Journal of Psychiatry*, *159*, 1563–1569.
- Hojat, M., & LaNoue, M. (2014). Exploration and confirmation of the latent variable of the Jefferson Scale of Empathy. *International Journal of Medical Education*, *5*, 73–81.
- Hojat, M., Mangione, S., Nasca, T. J., Cohen, M. J. M., Gonnella, J. S., Erdmann, J. B., & Magee, M. (2001). The Jefferson Scale of Physician Empathy: Development and preliminary psychometric data. *Educational and Psychological Measurement*, *61*(2), 349–365.
- Hojat, M., Vergare, M. J., Maxwell, K., Brainard, G., Herrine, S. K., Isenberg, G. A., Veloski, J., & Gonnella, J. S. (2009). The devil is in the third year: A longitudinal study of erosion of empathy in medical school. *Academic Medicine*, *84*(9), 1182–1191.
- Hollan, D. (2012). Emerging issues in the cross-cultural study of empathy. *Emotion Review*, *4*(1), 70–78.
- Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1–55.
- Kataoka, H. U., Koide, N., Ochi, K., Hojat, M., & Gonnella, J. S. (2009). Measurement of empathy among Japanese medical students: Psychometrics and score differences by gender and level of medical education. *Academic Medicine*, *84*(9), 1192–1197.
- Koopman-Holm, B., & Tsai, J. L. (2017). The cultural shaping of compassion. In E. M. Seppälä, E. SimonThomas, S. L. Brown, M. C. Worline, C. D. Cameron, & J. R. Doty (Eds.), *The Oxford handbook of compassion science* (pp. 273–285). Oxford University Press.
- Lawrence, E. J., Shaw, P., Baker, D., Baron-Cohen, S., & David, A. S. (2004). Measuring empathy: Reliability and validity of the empathy quotient. *Psychology in Medicine*, *34*(5), 911–919.
- Paro, H. B. M. S., Daud-Gallotti, R. M., Tibério, I. C., Pinto, R. M. C., & Martins, M. A. (2012). Brazilian version of the Jefferson Scale of Empathy: Psychometric properties and factor analysis. *Medical Education*, *12*, 73–79.
- Piumatti, G., Abbiati, M., Baroffio, A., & Gerbase, M. W. (2020). Empathy trajectories throughout medical school: Relationships with personality and motives for studying medicine. *Advances in Health Sciences Education*, *25*, 1227–1242.
- Ponnampuruma, G., Yeo, S. P., & Samarasekera, D. D. (2019). Is empathy change in medical school geo-socioculturally influenced? *Medical Education*, *53*, 655–665.
- Preusche, I., & Wagner-Menghin, M. (2013). Rising the challenge: Cross-cultural adaptation and psychometric evaluation of the

- adapted German version of the Jefferson scale of physician empathy for students (JSPE-s). *Advances in Health Sciences Education in Theory and Practice*, 18, 573–578.
- Quince, T., Parker, R. A., Wood, D. F., & Benson, J. A. (2011). Stability of empathy among undergraduate medical students: A longitudinal study at one UK medical school. *BMC Medical Education*, 11(1), 90–99.
- Quince, T., Thiemann, P., Benson, J., & Hyde, S. (2016). Undergraduate medical students' empathy: Current perspectives. *Advances in Medical Education and Practice*, 7, 443–455.
- Roff, S. (2015). Reconsidering the “decline” of medical student empathy as reported in studies using the Jefferson Scale of Physician Empathy-student version (JSPE-S). *Medical Teacher*, 37(8), 1–4.
- Silke, C., Brady, B., Boylan, C., & Dolan, P. (2018). Factors influencing the development of empathy and pro-social behavior among adolescents: A systematic review. *Children and Youth Services Review*, 94, 421–436.
- Smith, K. E., Norman, G. J., & Decety, J. (2017). The complexity of empathy during medical school training: Evidence for positive changes. *Medical Education*, 51, 1146–1159.
- Spasenoska, M., Costello, S., & Williams, B. (2016). Investigating the psychometric properties of the Jefferson scale of physician empathy in a sample of Malaysian medical students. *Advances in Medical Education and Practice*, 7, 331–339.
- Stanfield, R. B., Schwartz, A., O'Brien, C. L., Dekhtyar, M., Dunham, L., & Quirk, M. (2016). Development of a metacognitive effort construct of empathy during clinical training: A longitudinal study of the factor structure of the Jefferson scale of empathy. *Advances in Health Sciences Education Theory and Practice*, 21, 5–17.
- Szanto, T., & Krueger, J. (2019). Introduction: Empathy, shared emotions, and social identity. *Topoi*, 38, 153–162.
- Tavakol, S., Dennick, R., & Tavakol, M. (2011). Psychometric properties and confirmatory factor analysis of the Jefferson scale of physician empathy. *BMC Medical Education*, 11, 54–58.
- Taylor, Z. E., Eisenberg, N., Spinrad, T. L., Eggum, N. D., & Sulik, M. J. (2013). The relations of ego-resiliency and emotion socialization to the development of empathy and prosocial behavior across early childhood. *Emotion*, 13(5), 822–831.
- Vallabh, K. (2011). Psychometrics of the student version of the Jefferson Scale of Physician Empathy (JSPE-S) in final-year medical students in Johannesburg in 2008. *South African Journal of Bioethics Law*, 4(2), 63–68.
- Williams, B., & Beovich, B. (2019). A systematic review of psychometric assessment of the Jefferson Scale of Empathy using the COSMIN risk of bias checklist. *Journal of Evaluation in Clinical Practice*, 1, 14. <https://doi:10.1111/jep.13293>
- Williams, B., Brown, T., Boyle, M., & Dousek, S. (2013). Psychometric testing of the Jefferson scale of empathy health profession students' version with Australian paramedic students. *Nursing in Health Sciences*, 15, 45–50.
- Zenasni, F., Boujut, E., Bluffel du Vaure, C., Catu-Pinault, A., Tavani, J. L., Rigal, L., Jaury, P., Magnier, A. M., Falcoff, H., & Sultan, S. (2012). Development of a French-language version of the Jefferson Scale of Physician Empathy and association with practice characteristics and burnout in a sample of general practitioners. *International Journal of Person Centered Medicine*, 2(4), 759–766.