Psychometry and Cut-off Points of the Jefferson Scale of Empathy in Medical Students in Central America and the Caribbean

Víctor P. Díaz-Narváez*, Aracelis Calzadilla-Núñez†, Alejandro Reyes-Reyes‡, María G. Silva-Vetri¶, Pilar Torres-Martínez§, Eugenia Smirna González-Díaz**, Luz Marina Alonso-Palacio††, Jorge Luis Bilbao-Ramírez‡‡, Nuvia Estrada-Méndez¶¶, Dilia Esther Aparicio§§, Elizabeth Fajardo-Ramos***, Juan Carlos Ferrero-Otero***, Douglas Barrera-Gil†††, Ana Liliana Ríos-García‡‡‡

Objective: The objective of this study was to determine cut-off points that can be used to differentiate measures of empathy, which would then be classified as high, medium, or low. To do so, we used data from students from 7 medical schools in Colombia, El Salvador, and the Dominican Republic, after determining the psychometric properties of the 3-dimensional model of empathy in the Jefferson Scale of Empathy, S-version (for medical students).

Material and Methods: This non-experimental descriptive study had a sample that consisted of 6291 students. The structure and factor invariance were analyzed by country and sex. A hierarchical cluster analysis and a bifactorial analysis of variance were applied.

Results: The measure of empathy was reliable on the global scale (α = .82; ω = .88). A confirmatory factor analysis showed that the original model was replicable and adjusted to the data (comparative fit index [CFI] = .90; goodness of fit index = .94), while the multigroup analysis allowed to assume an invariant factor structure by country and gender (ΔCFI < .01). Tables were constructed with cut-off points for empathy and its dimensions.

Discussion and Conclusion: Our study solves the problem of comparing the scores and the levels of empathy observed in the medical students at different schools of medicine, making said comparisons within and between countries and between genders. The instrument used has adequate psychometric properties and the cut-off values obtained allow the classifying of people with lower or higher levels of empathy.

Key words: Empathy, Psychometrics, Medical student, Central America, Caribbean region

Medicine is a profession in which the practitioner must use interpersonal skills to achieve an optimal doctor-patient relationship (1). Empathy is an attribute made up of cognitive and emotional components that allows the practitioner to understand and communicate that understanding, increasing the quality of the relationship (2,3). The nature of this attribute is complex; the teaching/learning of empathy must be built implicitly and explicitly into the entire training process (4–9). It is suggested that empathy can be an object of learning, and the university is the last window for its development since the neurobiological bases that support this attribute reach their final development at around the age of 25 years (7–11). Knowing this, one of the natural interests of every university should consist of carrying out interventions to improve empathy in students. Most of the students in Latin America begin their professional practices in health services in primary care settings, all of which involve massive contact with patients. Due to the particular characteristics of this care setting, it is very important that these students begin their professional work with the greatest possible empathic development.

*Dental faculty, Universidad Andres Bello, Santiago, Chile; †Faculty of Health Sciences, Universidad Bernardo OHiggins, Santiago, Chile; ‡Faculty of Social Sciences and Communications, Universidad Santo Tomás, Concepción, Chile; ¶Dental faculty, Universidad Nacional Pedro Henríquez Ureña, Santo Domingo, Dominican Republic; §Dental faculty, Universidad San Sebastián, Concepción, Chile; **Faculty of Medicine, Universidad Central del Este, San Pedro de Macorís, Dominican Republic; ††Faculty Division of Health Sciences, Universidad del Norte, Barranquilla. Colombia; ‡‡Faculty of Medicine, Universidad Metropolitana, Barranquilla, Colombia; ¶¶Dental faculty, Universidad Evangélica de El Salvador, San Salvador, El Salvador; §§Faculty of Medicine, Corporación Universitaria Rafael Núñez, Cartagena, Colombia; ***Faculty of Health Sciences, Universidad del Tolima, Ibagué, Colombia; †††Faculty of Health Sciences, Universidad Andres Bello, San Salvador, El Salvador; ‡‡‡Faculty Division of Health Sciences, Universidad del Norte, Barranquilla, Colombia

The authors have no conflict of interest to disclose.

Address correspondence to: Dr. Víctor Patricio Díaz-Narváez, PhD. Dirección de Investigación, Universidad Andres Bello, Calle Echaurren 237, Santiago Centro, Región Metropolitana, Chile. Email: vicpadina@gmail.com; victor.diaz@unab.cl

One of the most commonly used instruments to measure empathy in Latin America is the Jefferson Scale of Empathy, S-version (JSE-S), for medical students. This instrument measures 3 components or dimensions: compassionate care (CC), perspective taking (PT), and walking in the patient's shoes (WIPS) (8,9,12–15). It has stability (supports construct validity) and acceptable reliability (Cronbach's alpha greater than .80).

Not all empathy research presents a psychometric study before analyzing the distribution of empathy levels. However, doing so is especially necessary because the parameters related to the estimation of the confirmatory factor analysis (CFA) can be affected by sample sizes smaller than 200. Currently there are no studies that estimate the relevant cut-off points. Knowledge of these points would have immediate practical utility, above all, to establish comparative results before and after the intervention, among other possible types of comparisons (1,3,5–14).

Using a sample group of students from 7 medical schools spread across 3 countries (Columbia, El Salvador, and the Dominican Republic), we aimed to define cut-off values for each of the 3 dimensions of the JSE-S (after determining its psychometric properties), enabling us to classify each participant as having a high, medium, or low level of empathy.

Materials and Methods

This is a descriptive, a posteriori, non-experimental study. Our sample was made up of secondary data from 7 schools of medicine in Colombia, the Dominican Republic, and El Salvador, collected from the years of 2016 through 2019. The sample studied consisted of 6291 (n) out of a total of 7729 (N) students (81.39% of the total).

Our research was approved by the Ethics Committee of Universidad San Sebastián (resolution 2015-02 and 2020-02).

Data analysis

The dependent variables were empathy levels, CC levels, PT levels, WIPS levels, and the number of clusters, while the independent variables were country and gender.

The sample size obtained was made up of all those students who voluntarily responded to the instrument. Those who did not respond (18.61%) may have been absent at the time the survey was deployed, may not have wished to participate, or were unable to do so (e.g., ill)

We subjected the primary data on empathy from all the medical schools studied to normality and homoscedasticity tests using the Kolmogorov-Smirnov and Levene's tests, respectively. We evaluated internal consistency by Cronbach's α, intraclass correlation coefficient, and McDonald's Omega coefficient (16,17).

To analyze the factorial structure of the JSE-S, we used the CFA (with oblique rotation) for each sample by country and gender, with a maximum likelihood estimation (MLE) based on Pearson's covariance matrix (18); and we examined factorial invariance by a multigroup analysis model, using as measures of goodness of fit the chi-square index ($\chi 2$), the normed chi-square ($\chi 2/df$), the goodness-of-fit index (GFI), the comparative fit index (CFI), the root mean square error of approximation (RMSEA), and standardized root mean square residual (SRMR) (19). We rated indicators that were approximately .90 as acceptable and those that were approximately .95 as excellent, in terms of the CFI, GFI, and AGFI (adjusted GFI), and we judged values of approximately .08 as acceptable settings and of approximately .06 as excellent settings, in terms of the RMSEA and SRMR (20–22). We considered factor loads greater than or equal to .40 as being significant (23).

Subsequently, we applied a hierarchical cluster analysis for posterior cases. We performed the clustering using the group-linking method (centroid grouping) and interval measurement, while we estimated the distance between clusters using the squared Euclidean method. Using the following, we determined several means within each cluster to see whether there was consistency between the robust means: Huber's M-estimator, Tukey's Biweight, Hampel's M-estimator, and Andrew's wave. We also compared their values with the arithmetic mean used in our study. In addition, we estimated the standard deviation of each cluster.

We conducted the comparison between means using a bifactorial analysis of variance (ANOVA). We calculated the weighted means using the university factor, while we used educational level as a covariable. We calculated the effect size (eta-squared [$\acute{\eta}^2$]) and the power of the analysis (PA: 1–β). We estimated the value of the coefficient of determination (R^2) by university, using a weighted least squares regression.

We performed the normalization of the groups by classifying the students and doing that by estimating percentiles with standardized data values. We present the results in the tables below. We performed all the analyses with IBM SPSS Statistics 25 and Amos 27. An α less than .05 was considered significant, as was a β less than or equal to .20.

Results

The students who participated in this study came from the following countries (and studied at medical schools within those countries): the Dominican Republic (2 institutions: Universidad Central del Este [n = 1144] and Universidad Nacional Pedro Henríquez Ureña [n = 1194]), Colombia (4 institutions: Universidad del Norte in Barranquilla [n = 695], Universidad Metropolitana in Barranquilla [n = 1586], Corporación Universitaria Rafael Núñez in Cartagena [n = 756], and Universidad de Tolima in Ibagué [n = 281]), and El Salvador (1 institution: Universidad Evangélica del Salvador in San Salvador [n = 635]).

Of the 6291 students in the sample, 4157 (66.08%) were women and 2134 (33.92%) were men. The results of the normality and homoscedasticity tests were not significant ($P > 0.05$); therefore, the data on empathy and its components were distributed in a normal way and with equal variance.

We calculated the reliability of global empathy for the total sample, made up of the set of schools, observing adequate estimates for Cronbach's alpha (α = .819) and McDonald's Omega (ω = .876). The intraclass correlation coefficient was .819 (95% CI: .813-.826; F = 5.537; P < .001). Cronbach's α values in the dimensions were satisfactory in PT (α = .842) and CC (α = .748); there was poor reliability in the WIPS component (α = .512). The coefficients were consistent with the estimates for each country (Table 1).

Factor structure (CFA)

The CFA estimation—using the ML method—of the JSE yielded statistically significant estimates for all the items, with standardized factor loadings that varied from λ = .065 to λ = .811, but not all the items loaded adequately on their corresponding factors; we observed factorial loads of less than .40, in particular in item 18 (.065 ≤ λ ≤ .210), item 1 (.300 ≤ λ ≤ .415), and item 17 (.303 ≤ λ ≤ .420) (Table 2).

Table 1. Descriptive statistics by country and gender, and JSE-S reliability coefficients

Dimension		Female		Male		Total			
		Mean	SD	Mean	SD	Mean (95% CI)	SD	α	ω
Dominican Republic (n = 2338)	CC	39.70	9.49	37.44	10.53	39.15 (38.75-39.55)	9.80	.763	
	PT	59.99	8.95	58.34	10.45	59.59 (59.21-59.97)	9.36	.859	
	WIPS	7.94	2.98	7.50	2.88	7.83 (7.71-7.95)	2.96	.643	
	E	107.62	15.40	103.27	16.05	106.57 (105.94-107.21)	15.67	.805	.845
Colombia (n = 3318)	CC	40.19	9.35	38.70	10.14	39.66 (39.33-39.98)	9.66	.747	
	PT	58.83	9.20	56.91	10.03	58.14 (57.82-58.47)	9.55	.838	
	WIPS	7.78	2.54	7.69	2.72	7.75 (7.66-7.84)	2.60	.386	
	E	106.81	16.70	103.30	17.21	105.54 (104.97-106.12)	16.97	.832	.882
El Salvador (n = 635)	CC	36.92	8.62	45.00	7.29	41.70 (41.02-42.39)	8.80	.705	
	PT	60.24	7.28	60.24	9.18	60.24 (59.58-60.90)	8.45	.784	
	WIPS	6.93	2.93	8.04	2.83	7.59 (7.36-7.81)	2.92	.564	
	E	104.09	13.15	113.28	15.07	109.53 (108.36-110.70)	15.01	.787	.824
Total (n = 6291)	CC	39.78	9.40	39.48	10.15	39.68 (39.44-39.91)	9.66	.748	
	PT	59.41	9.01	57.88	10.07	58.89 (58.66-59.12)	9.41	.842	
	WIPS	7.79	2.77	7.70	2.79	7.76 (7.69-7.83)	2.78	.512	
	E	106.99	15.97	105.05	16.98	106.33 (105.92-106.73)	16.35	.819	.876

CC: compassionate care; CI: confidence interval; E: empathy; PT: perspective taking; WIPS: walking in patient's shoes; α: Cronbach's alpha; ω: McDonald's Omega coefficient

Table 2. Standardized regression weights of the original JSE-S by country and gender

Factor	Item	Total sample	Colombia	Dominican Republic	El Salvador	Female	Male
PT	P2	.639	.537	.697	.634	.619	.659
	P4	.654	.775	.651	.638	.631	.694
	P5	.490	.494	.510	.367	.465	.528
	P9	.603	.593	.625	.410	.614	.577
	P10	.715	.649	.725	.751	.706	.723
	P13	.539	.475	.620	.480	.525	.564
	P15	.542	.596	.523	.452	.531	.564
	P16	.731	.811	.773	.690	.733	.733
	P17	.372	.345	.420	.303	.350	.407
	P20	.679	.616	.728	.619	.662	.701
CC	P1	.363	.325	.415	.300	.346	.412
	P7	.692	.617	.691	.676	.675	.720
	P8	.638	.648	.682	.598	.628	.653
	P11	.597	.545	.629	.651	.570	.646
	P12	.556	.443	.706	.609	.541	.584
	P14	.762	.774	.663	.669	.753	.781
	P18	.124	.210	.065	.082	.111	.154
	P19	.506	.608	.429	.379	.483	.547
WIPS	P3	.540	.491	.612	.524	.538	.547
	P6	.637	.487	.774	.752	.628	.653

CC: compassionate care; E: empathy; PT: perspective taking; WIPS: walking in patient's shoes. Bold indicates factor weights < .40

Based on the total sample, when relating the factors, we observed a correlation of .56 between PT and CC, .32 between CC and WIPS, and -.08 between PT and WIPS. The CFAs generated a similar pattern of results for the Dominican Republic and El Salvador, with adequate GFIs and with little adjustment being observed for the Colombian subsample. This confirmed the fit of the original 3-factor model, both in the total sample and segmented by country and gender (Table 3).

Analysis of invariance

We carried out a multigroup CFA, establishing an adequate base model by country (χ^2 = 5474.92; P = .0001; χ^2/df = 11.493; GFI = .924; CFI = .879; RMSEA = .042 [90% CI: .041-043]; SRMR = .043) and by gender (χ^2 = 3987.99; P = .0001; χ^2/df = 12.620; GFI = .940; CFI = .896; RMSEA = .043 [90% CI: .042-.044]; SRMR = .0499).

When examining the factorial invariance between countries and between genders, establishing nested models based on the base model, we observed significant changes in the chi-square value, which is reasonable, given the high sensitivity of this statistic to sample size (24). However, the differences in CFI are not relevant (ΔCFI ≤ .009, according to countries, and ΔCFI ≤ .003, according to gender), with all the values less than .01 indicating factorial invariance (25) (Table 4).

Table 3. Goodness-of-fit indices of the JSE-S CFA for the total sample and by country

CFA Model	χ^2	df	p	χ^2/df	GFI	AGFI	SRMR	CFI	RMSEA (90% CI)
Total sample	3680.797	158	.000	23.296	.944	.926	.047	.900	.060 (.058-.061)
Colombia	5008.014	163	.000	30.724	.861	.821	.069	.775	.095 (.092-.097)
El Salvador	350.660	164	.000	2.138	.948	.934	.046	.932	.042 (.036-.048)
Dominican Republic	869.619	162	.000	5.368	.965	.954	.045	.950	.043 (.040-.046)

AGFI: adjusted goodness of fit index; CFI: comparative fit index; CI: confidence interval; GFI: goodness of fit index; RMSEA: root mean square error of approximation; SRMR: standardized root mean square residual

Table 4. Goodness of fit of the multigroup confirmatory factor model according to gender and country in successive nested models

Model	χ^2	df	p	Δ χ^2	Δdf	p	CFI	ΔCFI
Invariance by gender								
Base model/ configural invariance	3987.990	316	.000	--	--	--	.896	--
Metric invariance	4031.106	333	.000	43.116	17	.000	.896	.000
Scalar invariance	4130.718	339	.000	99.612	6	.000	.893	.003
Invariance by country								
Base model/ configural invariance	5102.911	444	.000	--	--	--	.879	--
Metric invariance	5732.667	478	.000	629.756	34	.000	.870	.009
Scalar invariance	6047.609	490	.000	314.942	12	.000	.863	.007

CFI: comparative fit index

Table 5. Results of the sources of variation of the applied analysis of variance and estimation of the coefficient of determination of empathy and of each of its components

	Empathy (E)				Compassionate Care (CC)			
Source of variation	F	p	ή²	PA	F	p	ή²	PA
Course (covariable)	3.6	0.058	0.001	0.476	0.03	0.987	0.0005	0.05
Cluster	7595.3	0.0001	0.707	1.0	6502.5	0.0001	0.674	1.0
G	0.04	0.842	0.0005	0.055	0.121	0.728	0.0005	0.064
Cluster x G	10.14	0.0005	0.003	0.986	5.59	0.004	0.002	0.858
R^2 (a) (%)			71.4				68.4	
	Perspective Taking (PT)				Walking in Patient's Shoes (WIPS)			
Course (covariable)	3.01	0.083	0.0005	0.411	0.237	0.626	0.0005	0.078
Cluster	3584.3	0.0001	0.533	1.0	10461.1	0.0001	0.769	1.0
G	16.19	0.0005	0.003	0.98	0.097	0.341	0.0005	0.159
Cluster x G	6.18	0.002	0.002	0.893	1.87	0.155	0.001	0.391
R^2 (a) (%)			54.1				77.6	

F: Fisher's F test; G: gender; PA: power of the analysis or probability of committing type II error (1–β); p: probability of committing type I error associated with F statistic; ή²: eta-squared statistic; x: symbol of the interaction between factors

We present the results of the comparison between the clusters found within empathy and within its components in Table 5. The course covariate was not significant in any of the dimensions ($P > .05$).

The gender factor was significant in the PT dimension ($P = .005$), and the interaction between cluster and gender was significant in empathy (p = .0005) and in the CC and PT dimensions ($P = .004$ and P = .002, respectively). The previously described differences for the PT dimension in the gender factor and the interactions described are interesting findings, but they do not belong to the objectives of our manuscript and should be studied in later works. The differences between the clusters in empathy and in all its dimensions were highly significant ($P = .0001$), and the values of the effect size and the PA (1-β) in this factor were highly satisfactory.

Additionally, we can consider the values of the R^2 as high. Table 6 shows the normalized values in the clusters found in empathy and its components. The minimum and maximum values observed for each of the clusters can be considered as cut-off values to differentiate high, medium, and low values of empathy or in their respective dimensions, which could be useful in the qualitative classification of students from the point of view of the levels of empathy and the corresponding levels in each of the dimensions. The values of the robust estimators of the means, estimated in each of the clusters associated with empathy and to each of its dimensions, were all (without exception) within the confidence interval of the arithmetic mean estimated for the corresponding clusters.

Discussion and Conclusion

The total mean reached by medical students was 106.33 (SD = 16.35). Based on the results we observed, this can be classified as high, though below the 5th percentile. In terms of the CC and AP dimensions, it was also classified as high but was included in the 10–25 and 25–50 percentiles, respectively, while the WIPS dimension could be classified as average (above the 95th percentile). However, the standard deviations of these measurements were relatively high and showed that there was a significant number of students who should be classified as having medium levels of empathy and its dimensions. The foregoing shows that ample room remains for research into the developing of empathy in medical students, in particular, in those from the countries and universities from

Table 6. Results of the estimation of the percentiles of empathy and its components in each observed cluster and their corresponding minimum and maximum values

			Percentiles								
Empathy	Cluster	n	5	10	25	50	75	90	95	Min.	Max.
E	1	3415	108	109.5	113	118	124	129	132	107	140
	2	2860	77	79	84	93	101	105	105	56	106
	3	16	32	32	33	43	50	51	53	32	55
CC	1	4751	35	36	40	44	49	52	54	33	56
	2	1310	22	23	26	29	30	32	32	21	32
	3	230	8	11	13	16	19	20	20	8	20
PT	1	5775	49	51	56	62	66	69	70	45	70
	2	101	10	10	16	23	28	31	32	10	32
	3	415	35	37	40	41	43	44	44	33	44
WIPS	1	2879	8	8	8	9	10	11	12	8	12
	2	2973	2	3	5	6	7	7	7	2	7
	3	439	13	13	13	14	14	14	14	13	14

CC: compassionate care; E: empathy; Max.: maximum value; Min.: minimum value; PT: perspective taking; WIPS: walking in patient's shoes

which the sample population of this study was drawn. Moreover, our results underline the urgent need to take appropriate measures to ensure such development occurs. Furthermore, we observed that the total mean of the empathy scores was slightly lower than that of other samples of medical students from international studies, which generally is 112 (standard deviation of about 12) (26). There is a slightly higher mean in women than in men ($P < .001$), with a small effect size (d = .12), indicating that this difference, from a practical perspective, would not be significantly relevant (27,28).

The reliability of the measure, the Cronbach's alpha of .819, slightly exceeds the measures of reliability of estimates made in a diversity of samples of Latin American university students (range: .70 to .80, with a mean of .78), and with an intraclass correlation of .819, as well, which is also indicative of good reliability (26,29). It is possible to infer that the JSE-S is a reliable measure to be used in medical students in the countries studied and in Latin America, in general.

After having studied the factorial structure of the JSE-S, we developed a 3-factor model that fit the data well. In so doing, we confirmed the PT factor (which considered 10 items with factor loadings of .37 or more), observed low factor loading (<.40) for item 17, and found an α coefficient of .84. The CC factor included 8 items with factor loadings equal to or greater than .12, with a low factor load in items 1 and 18 and an α of .75. The third factor, putting oneself in the patient's shoes, included 2 items with factor loadings of .54 and .64, respectively, and a coefficient α of .51. This 3-factor model coincides with those reported in American, Spanish, and Turkish medical students (12,30,31). The low internal consistency of this last factor can be well explained by the low number of items that comprise it. To be stable, a factor structure requires a minimum of 3 or 4 items per factor (32,33).

Despite the low factorial load presented by 3 items on the scale (items 1, 17, and 18), their presence does not affect the reliability of the measure, but they have high error variance (with squared multiple correlations of .132, .139, and .015, respectively). While, in our study, we considered standardized factor loadings greater than .40 significant, a good general rule is that standardized factor loads should be greater than .50 or, ideally, greater than or equal to .70 (applying to items that, individually, do not contribute mainly to the measurement of the respective factors or to the general empathy-measurement model) (34).

The results of the ANOVA led us to observe that it was important to have studied the covariate, specifically in the PT dimension. Some factors and interactions were significant in the empathy component and in its dimensions, all of which would imply the existence of variability in the distribution those factors and interactions.

However, the only factor studied in which the effect size ($\eta 2$) and the PA values were high and very satisfactory (at the same time) in the clusters was factor WIPS, showed that the total variance depends more on the differences between the clusters than it does on the other factors and interactions. However, the significance found in some of the factors of and interactions in empathy and in some of its dimensions shows the existence of empathic variability, which has already been reported in other works (2,3,6,7,9,10). These findings are not related to the objectives of this article but should be studied in future works. The inclusion of these factors and of the interactions studied in the model allowed us to extract from the total error of the model all the variance attributable precisely to these factors and interactions. In other words, the interference of the errors of the factors and significant interactions on empathy and its dimensions (as the case may be) was eliminated from the total error of the model. We believe that the differences we found between the clusters and their magnitude are genuine and that these represent a good description of the studied phenomenon. The fact that the effect sizes were large and the PA in the cluster factor assumed the highest possible value (1.0) means not only that the existence of differences between the clusters has been verified but that these differences are large and that the probability of making a type II error is almost null.

We have not found any studies in the literature that have estimated cut-off points based on a previous confirmation of the Hojat model (1–16,30–39). The reliability and confirmation of the existence of equality of the latent 3-dimensional model in each of the regions studied, including with respect to genders, are relevant because they allow comparisons of the levels of empathy (and its dimensions) between the regions examined, and the same occurs with respect to gender. Another finding is that the hierarchical cluster analysis was able to consistently discriminate 3 clusters in construct empathy and in the 3

dimensions observed by using the arithmetic mean in each case. Consistent with these results, the estimated robust means coincide with the values of the arithmetic means used in the comparisons to determine differences between clusters (those estimated means were within the 95% confidence interval of the arithmetic mean). Adding consistency to these results were the R^2 values.

These values show that the factors and interactions studied explain the relatively high percentage of the distribution of empathy levels in the countries studied. Additionally, the largest source of the explained variance can be attributed to the differences between the clusters.

The estimations of the minimum and maximum values associated with each cluster could suggest that these values constitute a first approximation of what could be called "cut-off points" for high, medium, and low levels. Within each cluster, it was possible to differentiate the values of empathy and their dimensions associated with percentiles. These percentiles allowed qualitative discrimination within each of the 3 levels mentioned above: high, medium, and low.

Several current published studies have attempted to prove that certain types of interventions might represent agents of change in terms of measuring empathy, most particularly using the JSE-S in medical students or other students in the health sciences. However, the results of these studies measure the changes using statistical tests that could produce significant differences, which differences cannot classify the level of success obtained without reference points that measure the quality of the intervention.

Given that the concept of empathy is complex, the construction of cut-off points becomes important to assess the scores and rate the change in empathy levels. Achieving benchmarks that can demonstrate the true meaning and impact of an intervention (3,6,8–15). As a consequence, our results have an immediate, practical application.

Study limitations

The values found in our study, which values allow the classification of students as having higher or lower levels of empathy (and their components), may be valid for medical students belonging to the regions studied (South America, Central America, and the Caribbean), but they cannot be extended to regions not examined in the present research.

We conclude that the JSE-S has adequate psychometric properties in samples of medical students from South America, Central America, and the Caribbean, resulting in a reliable and valid measure that allows the establishing of useful cut-off values to classify medical students with lower or higher levels of empathy (and its components). Therefore, these results solve the problem of comparing the observed results of the levels of empathy between medical schools within and between countries and pre- and post-intervention since such comparisons were only possible using traditional statistical methods. Consequently, these results have immediate practical application.

Resumen

Objetivo: El objetivo de este estudio fue determinar puntos de corte que puedan utilizarse para diferenciar medidas de empatía que luego se clasificarían como: alto, medio y bajo. Para ello utilizamos datos de estudiantes de 7 facultades de medicina de Colombia, El Salvador y República Dominicana, previa determinación de las propiedades psicométricas del modelo tridimensional de empatía en la Escala Jefferson de Empatía, versión S (para estudiantes de medicina). Material y Métodos: Este estudio descriptivo no experimental contó con una muestra de 6291 estudiantes. La estructura y la invarianza factorial fueron analizadas por país y sexo. Aplicamos un análisis de conglomerados jerárquico y un análisis bifactorial de varianza. Resultados: La medida de la empatía es confiable en la escala global (α = .82, ω = .88). El Análisis Factorial Confirmatorio mostró que el modelo original era replicable y ajustado a los datos (índice de ajuste comparativo [Comparative Fit Index, CFI] = .90; índice de bondad de ajuste = .94), mientras que el análisis multigrupo permitió asumir una estructura factorial invariante por país y género (ΔCFI < .01). Se construyeron tablas con puntos de corte para la empatía y sus dimensiones. Discusión y Conclusión: Nuestro estudio resuelve el problema de comparar las puntuaciones y los niveles de empatía observados entre las Facultades de Medicina dentro y entre países y entre géneros. El instrumento utilizado tiene adecuadas propiedades psicométricas y los valores de corte obtenidos permiten clasificar a las personas con menor o mayor nivel de empatía.

References

1. Santiago LM, Moreira MC, Silva I. Carta ao Editor acerca do artigo "Reconsiderando o 'Declínio' da Empatia com o Curso em Estudantes de Odontologia na América Latina" [Letter to the Editor about the article "Reconsidering the 'Decline' of Dental Student Empathy within the Course in Latin America"]. Acta Med Port. 2018;31(5):290-291. doi:10.20344/amp.10663
2. Díaz-Narváez VP, Alonso-Palacio LM, Caro SE, et al. Compassionate Care Component of the Construct Empathy in Medical Students in Colombia and Dominican Republic. Acta Med Mediterr. 2017;33:115-121. doi:10.19193/0393-6384_2017_1_018
3. Bilbao JL, Iglesias JE, Díaz-Narváez VP, Crespo-Camacho E, Alonso LM, Alcocer A. Orientación empática en estudiantes de medicina de las Universidades Libre y San Martin, Barranquilla, Colombia. Rev Fac Med. 2015;63:657-663. doi:http://dx.doi.org/10.15446/revfacmed.v63.n4.49715
4. Melchers M, Montag C, Reuter M, Spinath FM, Hahn E. How heritable is empathy? Differential effects of measurement and subcomponents. Motiv Emot. 2016;40:720-730. https://doi.org/10.1007/s11031-016-9573-7
5. Díaz-Narváez VP, Estrada-Méndez N, Arévalo-López Y, Calzadilla-Núñez A, Utsman-Abarca R. Empathy, Components of Empathy and Curricular Evaluation of the Faculty of Dentistry, Evangelical University of El Salvador. Pesqui Bras Odontopediatria Clin Integr. 2017;17:e3759. http://dx.doi.org/10.4034/PBOCI.2017.171.40
6. Calzadilla-Núñez A, Díaz-Narváez VP, Dávila-Pontón Y, et al. Erosion of empathy during medical training by gender. A cross-sectional study. Erosión empática durante la formación médica según el género. Estudio transversal. Arch Argent Pediatr. 2017;115(6):556-561. doi:10.5546/aap.2017.eng.556

7. Díaz-Narváez VP, Bilbao JL, Estrada-Méndez N, Ulloque J, Calzadilla-Núñez A, Barrera-Gil D. Consideraciones acerca de la declinación empática en estudiantes de medicina en Latinoamérica. Rev Cienc Salud. 2020;18:1-16. https://doi.org/10.12804/revistas.urosario.edu.co/revsalud/a.9258
8. Mahoney S, Sladek RM, Neild T. A longitudinal study of empathy in preclinical and clinical medical students and clinical supervisors. BMC Med Educ. 2016;16(1):270. Published 2016 Oct 18. doi:10.1186/s12909-016-0777-z
9. Aparicio DE, Ramos AE, Mendoza J, Utsman-Abarca R, Calzadilla-Núñez A, Díaz-Narváez VP. Levels of empathy, empathy decline and differences between genders in medical students of Cartagena (Colombia). Educ Medica. 2019;20:136-143. https://doi.org/10.1016/j.edumed.2018.06.004
10. Barrera-Gil D, Estrada-Méndez N, Arévalo Y, Calzadilla-Núñez A, Díaz-Narváez VP. Empatía en estudiantes de medicina de la República de El Salvador: estudio transversal [Empathy in medical students in the Republic of El Salvador: Cross-sectional study]. J Healthc Qual Res. 2018;33(3):136-143. doi:10.1016/j.jhqr.2018.03.002
11. Tariq N, Rasheed T, Tavakol M. A Quantitative Study of Empathy in Pakistani Medical Students: A Multicentered Approach. J Prim Care Community Health. 2017;8(4):294-299. doi:10.1177/2150131917716233
12. Hojat M, DeSantis J, Shannon SC, et al. The Jefferson Scale of Empathy: a nationwide study of measurement properties, underlying components, latent variable structure, and national norms in medical students. Adv Health Sci Educ Theory Pract. 2018;23(5):899-920. doi:10.1007/s10459-018-9839-9
13. Kalyan SV, Manjula S, Padma TM, Pratap KV, Vineela P, Varma SC. Assessment of empathy among clinical dental students in a teaching dental institution in Telangana State, India. J Indian Assoc Public Health Dent. 2017;15:162-165. doi:10.4103/jiaphd.jiaphd_174_16
14. Williams B, Beovich B. Psychometric properties of the Jefferson Scale of Empathy: a COSMIN systematic review protocol. Syst Rev. 2019;8(1):319. Published 2019 Dec 10. doi:10.1186/s13643-019-1240-0
15. Hsiao CY, Tsai YF, Kao YC. Psychometric properties of a Chinese version of the Jefferson Scale of Empathy-Health Profession Students. J Psychiatr Ment Health Nurs. 2013;20(10):866-873. doi:10.1111/jpm.12024
16. Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika. 1951;16:297-334. https://doi.org/10.1007/BF02310555
17. McDonald RP. Test theory. A unified treatment. Psychology Press; 1999. https://doi.org/10.4324/9781410601087
18. Curran PJ, West SG, Finch JF. The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. Psychol Methods. 1996;1:16-29. https://doi.org/10.1037/1082-989X.1.1.16
19. Jöreskog KG. Simultaneous factor analysis in several populations. Psychometrika. 1971;36:409-426. http://files.eric.ed.gov/fulltext/ED055110.pdf
20. Bentler PM, Bonett DG. Significance tests and goodness of fit in the analysis of covariance structures. Psychol Bull. 1980;88:588-606. https://doi.org/10.1037/0033-2909.88.3.588
21. Browne MW, Cudeck R. Alternative ways of assessing model fit. Sociol Methods Res 1992;21:230-258. https://doi.org/10.1177/0049124192021002005
22. Kline RB. Principles and practice of structural equation modeling. 4th ed. The Guilford Press; 2015.
23. Stevens JP. Applied multivariate statistics for the social sciences. 5th ed. Routledge; 2009.
24. Lévy JP, Iglesias, V. Invarianza causal con muestras múltiples. In: Lévy JP, Varela J, eds. Modelización con Estructuras de Covarianzas en Ciencias Sociales. Netbiblo; 2006:279-318.
25. Cheung GW, Rensvold RB. Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance. Struct Equ Modeling. 2002;9:233-255. https://doi.org/10.1207/S15328007SEM0902_5
26. Hojat M. The Jefferson Scale of Empathy. In: Empathy in health professions education and patient care. Springer International; 2016:83-128.
27. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Routledge; 1988.
28. Cohen J. A power primer. Psychol Bull. 1992;112(1):155-159. doi:10.1037//0033-2909.112.1.155
29. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research [published correction appears in J Chiropr Med. 2017 Dec;16(4):346]. J Chiropr Med. 2016;15(2):155-163. doi:10.1016/j.jcm.2016.02.012
30. Ferreira-Valente A, Costa P, Elorduy M, Virumbrales M, Costa MJ, Palés J. Psychometric properties of the Spanish version of the Jefferson Scale of Empathy: making sense of the total score through a second order confirmatory factor analysis. BMC Med Educ. 2016;16(1):242. Published 2016 Sep 19. doi:10.1186/s12909-016-0763-5
31. Bilgel N, Ozcakir A. Turkish Version of the Jefferson Scale of Empathy Psychometric Properties. Eur Sci J. 2017;13:101-111. https://doi.org/10.19044/esj.2017.v13n20p101
32. Velicer WF, Fava JL. Affects [sic] of variable and subject sampling on factor pattern recovery. Psychol Methods. 1998; 3:231-251. https://doi.org/10.1037/1082-989X.3.2.231
33. Ferrando PJ, Anguiano-Carrasco C. El análisis factorial como técnica de investigación en psicología. Papeles Psicol. 2010;31:18-33.
34. Hair JF, Black WC, Babin BJ, Anderson RE. Multivariate data analysis. 7th ed. Pearson Education, Inc.; 2013
35. Díaz-Narváez VP, Erazo Coronado AM, Bilbao JL, et al. Reconsidering the 'Decline' of Dental Student Empathy within the Course in Latin America. Acta Med Port. 2017;30(11):775-782. doi:10.20344/amp.8681
36. Ye X, Guo H, Xu Z, Xiao H. Empathy variation of undergraduate medical students after early clinical contact: a cross-sectional study in China. BMJ Open. 2020;10(7):e035690. Published 2020 Jul 19. doi:10.1136/bmjopen-2019-035690
37. Yuguero O, Marsal JR, Esquerda M, Galvan L, Soler-González J. Cross-sectional study of the association between empathy and burnout and drug prescribing quality in primary care. Prim Health Care Res Dev. 2019;20:e145. Published 2019 Oct 30. doi:10.1017/S1463423619000793
38. Nathwani SJ, Vedd N. Medical students' perspective: project-based learning approach to increase medical student empathy. Med Educ Online. 2020;25(1):1794342. doi:10.1080/10872981.2020.1794342
39. Montanari P, Petrucci C, Russo S, Murray I, Dimonte V, Lancia L. Psychometric properties of the Jefferson Scale of Empathy-Health Professional Student's version: An Italian validation study with nursing students. Nurs Health Sci. 2015;17(4):483-491. doi:10.1111/nhs.12221