Psychometry: Cutting-Off Points and Standardization of the Jefferson Empathy Scale Adapted for Students of Kinesiology

Alejandro Reyes-Reyes¹, Aracelis Calzadilla-Núñez²,³, Pilar Torres-Martínez⁴, Patricia Díaz-Calzadilla³, Wilson Pastén-Hidalgo⁵, Fanny Bracho-Milic⁶, and Víctor Díaz-Narváez⁷

Abstract
Currently, the most common measurement of empathy is obtained using scales that offer a continuum between a minimum and a maximum value. The objectives of this study were to establish a norm and estimate cut-off points that would make it possible to assess the Jefferson Scale of Empathy (JSE) version for Health Professions students (HPS-version), and to determine its psychometric properties in Chilean physical therapy students. A secondary analysis was done on a data set from three schools of physical therapy (n = 850, 412 women [48.5%], and 438 men [51.5%]), applying confirmatory factor analysis (CFA) and hierarchical cluster analysis. A CFA replicated the original three-factor model of empathy with sufficiently fit the data. A hierarchical cluster analysis yielded four categories for the level of empathy: high, medium-high, medium-low, and low. Multi-group analyses supported the assumption of a gender-invariant factor structure. Results confirmed the reliability of the global scale (α = .835), and the Perspective Taking (α = .732), Compassionate Care (α = .842), and Walking in Patient’s Shoes (α = .686) dimensions. The instrument made it possible to establish four ordinal categories in the level of students’ empathy. We conclude that the HPS-version of the JSE has adequate psychometric properties; namely validity, reliability, and cut-off points that justify administering it to Chilean physical therapy students.

Keywords
empathy, perspective taking, compassionate care, walking in patient’s shoes, cut-off points

Introduction
Empathy is defined as the ability to understand feelings and emotions objectively and rationally; experiencing what other people feel and think (Díaz-Narváez & Calzadilla-Núñez, 2019; Díaz-Narváez et al., 2017; Preusche & Lamm, 2016; Svenaeus, 2016). Empathy is an attribute that plays an important role in the interaction between physical therapists and their patients. It involves both emotional and cognitive factors (Svenaeus, 2016), and is conditioned by their interactions (Díaz-Narváez et al., 2017). Some authors have argued that the cognitive component of empathy may be taught via disciplinary training processes, but the same does not occur with affective empathy (Díaz-Narváez et al., 2017; Preusche & Lamm, 2016). The latter seems to consolidate itself from the first processes of ontogeny and continues its formation until later adolescence (Díaz-Narváez et al., 2017). In this sense, the full integration of the longitudinal teaching of empathy in undergraduate curricula is important, and will positively influence patient care (Díaz-Narváez & Calzadilla-Núñez, 2019).

Studies of empathy (E) estimation in health-science students raise the existence of three components of this attribute (Ávila et al., 2020; Calzadilla-Núñez et al., 2017; Díaz-Narváez et al., 2015, 2017; Galán et al., 2014; González-Martínez et al., 2018; Lee & Seomun, 2016; Pastén-Hidalgo et al., 2019; Rueckert et al., 2011): the ability to feel compassion (Compassionate Care, CC), the ability to adopt the patient’s perspective (Perspective Taking, PT), and the ability to understand others (Walk in Patient’s Shoes, WIPS).
Compassionate care is associated with the emotions of the subject, and seems to be influenced by both biology (i.e., a product of evolution, ontogeny, and their interaction) and culture, and is associated with moral, altruistic, and religious behavior, among others (Calzadilla-Núñez et al., 2017; Díaz-Narváez et al., 2017). Perspective Taking (PT) is associated with a person’s ability to differentiate them-self from others (i.e., patients) and avoid “emotional contagion.”

The ability to understand others, or “walk in a patient’s shoes” (WIPS) refers to the ability to actively observe a subject to thereby penetrate their thinking. CC is part of the emotional component, while PT and WIPS are parts of the cognitive component (Calzadilla-Núñez et al., 2017).

There is a positive correlation between empathy and compassion (Lee & Seomun, 2016). People suffer from burnout, stress, excessive academic load, and compassion fatigue (among other conditions) are observed to have lower empathy (Hunt et al., 2017). These findings hint at the complex and multidimensional nature of empathy (Díaz-Narváez et al., 2017).

Regarding gender, while women are often perceived as more empathetic than men, both in the general population and among health students (Rueckert et al., 2011), results of empirical measurements of empathy in students from Latin America have shown contradictory results (Calzadilla-Núñez et al., 2017; Díaz-Narváez et al., 2017; González-Martínez et al., 2018).

Teaching empathy in Chilean universities is not formalized in their curricula (Díaz-Narváez et al., 2015). There is a trend of conducting “empathic interventions,” however this lacks prior definitions of this attribute’s behavior. Nevertheless, some authors recognize the possible errors of these activities and postulate that interventions of this type should involve modification of the entire curriculum to achieve permanent positive changes regarding empathy in students (Calzadilla-Núñez et al., 2017; Díaz-Narváez & Calzadilla-Núñez, 2019; Díaz-Narváez et al., 2015, 2017; Galán et al., 2014; González-Martínez et al., 2018; Madera-Anaya et al., 2016; Preusche & Lamm, 2016). As a consequence, an objective diagnosis of specific characteristics of empathic behavior is both important and necessary so that our interventions correspond to reality. However, currently no cut-off points allow us to divide students by classification such as “high,” “medium,” or “low.” Such classification is important, as it allows for pre- and post-intervention comparison within and between student groups from different schools and universities. The development of cut-off points requires that criteria for their delimitation be established, and clinical (or physiological) evidence for setting these values, or norms for generating them, be found; however, this does not currently exist, at least regarding empathy measurements produced with the JSE in Chile or in the rest of Latin America.

Norms for administering the JSE are recent, and nationwide guidelines only exist for US osteopathic medicine students (Hojat et al., 2018, 2019). In those studies conducted in Latin American countries (Chile, Colombia, Mexico, El Salvador, Ecuador, the Dominican Republic, Panama, and Puerto Rico, among others), JSE results have been reported using raw scores due to the absence of national or international norms, preventing researchers from adequately interpreting respondents’ performance relative to their reference populations.

A suitable adaptation of the JSE for Chilean students would require both developing a norm and setting cut-off scores to ensure the usefulness of the results obtained in research and empathy diagnosis. The aim of the present study is to establish such a norm and cut-off scores based on a criterion that distinguishes levels of empathy in physical therapy students in Chile, detailing the psychometric evidence in support of these constructs.

Methods

Participants

The participants were 850 physical therapy students, 412 women (48.5%), and 438 men (51.5%), selected via convenience sampling from University of Atacama (*n* = 191, 22.5%), Bernardo O’Higgins University (*n* = 484, 56.9%), and Universidad Mayor (*n* = 175, 20.6%), faculties located in Copiapó, Santiago, and Temuco. Respectively and covering the north, center, and south of Chile. The sample (*n* = 850) was randomly subdivided into two sub-samples. The first sub-sample (*n*₁ = 510, 60%) allowed establishing the cut-off scores that generate four levels of empathy. The second sub-sample (*n*₂ = 340, 40%) allowed us to classify the students according to the previously defined levels and to establish the norm. We compared the sub-samples using the Mann–Whitney *U* test, which revealed no statistically significant differences in empathy and its dimensions (all *p*-values > .05).

Instruments

The Jefferson Scale of Empathy (JSE) (Hojat et al., 2001; Hojat, Gonnella, Nasca, Mangione, Veloksi, et al., 2002) Spanish version for Health Professions Students (HPS-version) was administered. It is composed of 20 items, each with a 7-level Likert-type response (1 = strongly disagree, 7 = strongly agree), that make up a scale from 20 to 140 points. The instrument is intended to measure three components of empathy: Perspective Taking (10 items, 10–70 points), Compassionate Care (8 items, 8–56 points), and Walking in Patient’s Shoes (2 items, 2–14 points) (Hojat, Gonnella, Nasca, Mangione, Vergare, et al., 2002). The original scale and its multiple versions currently lack cut-off scores.

Procedures

This is a methodological study with a descriptive design, aimed at performing a secondary analysis of empathy data.
collected from 2015 to 2020. Before the application of the scale in its Spanish version, the cultural adaptation of the instrument was conducted using the criteria of raters as a procedure to adapt it to physical therapy students, and was applied to groups of between 30 and 50 students during regular class hours after they had given informed consent. This study was approved by the Ethics Committee of the Universidad San Sebastián, Chile (Resolution 2020-2).

Data Analysis

The data were subjected to normality testing via the Kolmogorov-Smirnov test (Kolmogorov, 1933). Using Levene’s (1960) test to evaluate the similarity of variables between sub-samples according to empathy categories. Reliability was evaluated in multiple ways: using Cronbach’s alpha (Cronbach, 1951) to establish internal consistency, the intraclass correlation coefficient (ICC) to establish the stability of data among universities (Mohamed & Shoukri, 2004), and McDonald’s coefficient omega (McDonald, 1999), which yields a more accurate measurement of reliability by considering the homogeneity of items through item-test correlation. To analyze the factor structure of the JSE, confirmatory factor analysis (CFA) was used, with maximum likelihood (ML) as recommended by Curran et al. (1996). To evaluate the fit of the models, various goodness of fit indices were considered: (a) chi-square index ($\chi^2$), a non-significant value indicated a good fit; (b) chi-square normed ($\chi^2/df$), considering values lower than 2 to indicate adequate adjustment; (c) goodness of fit index (GFI), comparative fit index (CFI), and adjusted goodness of fit index (AGFI), values $\geq .90$ indicated an acceptable fit and $\geq .95$ was indicative of a good fit; (d) Root mean square error of approximation (RMSEA), a value $\leq .05$ (90% CI $\leq .08$) was indicative of a good fit, and (e) standardized root mean square residual (SRMR), values around .08 were judged as acceptable fits, and around .06 as excellent fits (Bentler & Bonett, 1980; Browne & Cudeck, 1992; Hu & Bentler, 1999; Kline, 2005). Factor loadings greater than or equal to 0.40 were considered significant (Stevens, 1992).

Factor invariance was analyzed using a multi-group analysis model (Jöreskog, 1971), using the chi-square test ($\chi^2$) to assess goodness of fit, but since the chi square test is sensitive to the sample size, decreases in the CFI of less than .01 ($\Delta \leq .01$) in comparison with the previous model were considered to be the most adequate indicator of invariance (Cheung & Rensvold, 2002). To determine the cut-off scores, as no prior criteria existed, we conducted a hierarchical cluster analysis according to the recommendations of Hair et al. (2013). We employed a posteriori cases with standardized data and centroid clustering based on the squared Euclidean distance to the group together the cases of the first sample ($n_1 = 510$). This process yielded approximate values that made it possible to establish ranges with which to classify participants according to their empathy scale scores. Multiple statistics were calculated to describe the clusters, adding a Huber M-estimator due to the lack of symmetry in the distribution of the variables, which generated optimal values regardless of error distribution (Cajal et al., 2012). To determine whether sufficiently different clusters were generated, we estimated the difference between empathy means using a one-factor ANOVA and minimum significant difference method (MSD) to make multiple comparisons of the measurements, calculating effect size (partial eta squared: $\eta^2$), and the adjusted coefficient of determination ($R^2$). To validate the cut-off scores, we employed the second sample ($n_2 = 340$) and calculated its sensitivity and specificity relative to the overall scale score, to evaluate whether new cases had been accurately classified according to the cut-off scores established by the first sample ($n_1$). Finally, we established a norm that made it possible to estimate a percentile value based on scores yielded by the JSE and each of its dimensions according to two different levels of the measurement and according to the cut-off scores that we had set. The level of significance used was $\alpha < .05$ and $\beta \leq .20$. All analyses were conducted with IBM SPSS Statistics 25 and Amos 25 (Figure 1).

Results

Descriptive Statistics

The results of the normality and homoscedasticity tests were not significant ($p > .05$), therefore the data on empathy and its components were distributed normally and with equal variance. Descriptive statistics for the total sample and for the sample segmented by gender and university are presented in Table 1. The mean score of empathy for the total sample was 107.70 ($SD = 16.54$, range = 59–140); the mean score for men was 105.70 ($SD = 16.67$, range = 66–137) and for women it was 109.82 ($SD = 16.15$, range = 59–140), with statistically significant differences between the sub-samples ($t_{488} = 3.66$, $p < .0001$, $d = .251$).

Reliability Analysis

Reliability was estimated for the total sample, made up of the three universities, and reached suitable values for a global scale (Cronbach’s alpha, $\alpha = .835$, and McDonald’s Omega, $\omega = .832$). With satisfactory values in compassionate care ($\alpha = .842$), perspective taking ($\alpha = .732$) and walking in patient’s shoes ($\alpha = .686$) dimensions. Coefficients were consistent with the estimates made with each university’s sub-samples (see Table 1). The intraclass correlation coefficient was 0.835 (CI: 0.818; 0.851) and highly significant ($F = 6.05$, $p < .001$). When examining the homogeneity value of items by means of the corrected item-total correlation ($r$), a range was found between .095 and .699 with a median of .407. Seventy-five percent of the items presented an adequate value above 0.30, and five items (items 15, 17, 5, 18, and 20) displayed correlations below the expected value ($r_{15} = .095$, $r_{17} = .166$, $r_5 = .240$, $r_{18} = .278$, and $r_{20} = .279$).
Confirmatory Factor Analysis

To evaluate the construct validity of the scale, and confirm the structure of the latent variables of the JSE, a confirmatory factor analysis (CFA) was performed ($n_1 = 510$) that sought to test the theoretical structure of three factors and 20 items proposed by Hojat and his collaborators (Hojat et al., 2018). A baseline model with adequate adjustment has been established ($\chi^2 = 373.006, p = .0001; \chi^2/df = 2.275; GFI = .958; AGFI = .946; CFI = .951; \text{RMSEA} = .039 \ [90\% \text{ CI} = .034-.044]; \text{SRMR} = .043$), whose significant standardized factor loadings vary between $\lambda = .236$ and $\lambda = .790$ for the total sample. Models established by university and by gender present similar values.

The general model includes three items with factor loadings of below 0.40 (items 15, 17, and 18, see Table 2).

The CFA generates a similar pattern of results for each sub-samples, achieving an adequate adjustment to the global sample (made up of the three universities and divided by gender), with goodness-of-fit indices that confirm the adjustment of the original model of three factors to the samples studied (see Table 3).

Invariance Analysis

A factor invariance analysis was performed comparing women and men via a multigroup analysis. This analysis

**Figure 1.** General outline of the steps of the study for establishing the cut-off points and the JSE standard.
Table 1. Descriptive Statistics by University, Gender, and JSE Reliability Coefficients.

| University | Female (n=412) | Male (n=438) | Total (n=850) |
|------------|----------------|--------------|--------------|
|            | Mean | SD     | Mean | SD     | Mean [95% CI] | SD | α | ω |
| **UBO** (n=484) |     |        |      |        |              |    |   |   |
| CC         | 38.84 | 11.92  | 35.22 | 12.09  | 36.94 [35.86; 38.02] | 12.14 | .858 |
| PT         | 60.01  | 7.15   | 59.52 | 7.02   | 59.75 [59.12; 60.38] | 7.08  | .729 |
| WIPS       | 7.67   | 3.23   | 6.92  | 2.74   | 7.27 [7.01; 7.54] | 3.01  | .655 |
| E          | 106.52 | 17.13  | 101.65 | 16.74  | 103.97 [102.44;105.49] | 17.08 | .836 | .828 |
| **UM** (n=175) |     |        |      |        |              |    |   |   |
| CC         | 43.59  | 9.33   | 40.58 | 9.66   | 41.95 [40.52; 43.39] | 9.60  | .805 |
| PT         | 61.69  | 5.96   | 59.16 | 7.62   | 60.31 [59.27; 61.36] | 7.00  | .730 |
| WIPS       | 8.75   | 3.31   | 8.85  | 2.89   | 8.81 [8.35; 9.27] | 3.08  | .693 |
| E          | 114.03 | 13.68  | 108.59 | 15.91  | 111.07 [108.82;113.33] | 15.13 | .821 | .872 |
| **UDA** (n=191) |     |        |      |        |              |    |   |   |
| CC         | 42.94  | 9.25   | 43.82 | 8.49   | 43.45 [42.08;44.62] | 8.89  | .747 |
| PT         | 62.91  | 5.93   | 62.30 | 6.39   | 62.63 [61.75; 63.50] | 6.14  | .707 |
| WIPS       | 8.11   | 3.15   | 8.04  | 2.86   | 8.08 [7.65; 8.51] | 3.01  | .686 |
| E          | 113.96 | 13.94  | 114.17 | 13.26  | 114.06 [112.12; 116] | 13.59 | .785 | .860 |
| **Total** (n=850) |     |        |      |        |              |    |   |   |
| CC         | 40.78  | 11.04  | 38.13 | 11.49  | 39.41 [38.65; 40.18] | 11.34 | .842 |
| PT         | 61.06  | 6.74   | 60.00 | 7.12   | 60.51 [60.05; 60.98] | 6.95  | .732 |
| WIPS       | 7.99   | 3.25   | 7.57  | 2.91   | 7.77 [7.56; 7.98] | 3.08  | .686 |
| E          | 109.82 | 16.15  | 105.70 | 16.67  | 107.70 [106.58;108.81] | 16.54 | .835 | .832 |

Note. UBO=Universidad Bernardo O’Higgins; UM=Universidad Mayor; UDA=Universidad de Atacama; CC=compassionate care; PT=perspective taking; WIPS=walking in patient’s shoes; E=empathy; CI=confidence interval; α=Cronbach’s alpha; ω=McDonald’s Omega coefficient.

revealed a reasonably adequate, although not excellent, fit of the model to the data: $\chi^2 = 612.610, p < .0001$, $\chi^2/df = 1.868$, SRMR = .047, GFI = .933, AGFI = .915, CFI = .934, RMSEA = .032 (90% CI = 0.018–0.036).

Homologously, the invariance by university was analyzed: $\chi^2 = 849.490, p < .000$, $\chi^2/df = 1.727$, SRMR = .047, GFI = .933, AGFI = .915, CFI = .934, RMSEA = .032 (90% CI = 0.026–0.033). After establishing the baseline models by gender and university, nested models were established from the base model. Significant changes were observed in the chi-square value by university, which is reasonable given the high sensitivity of this statistic to the sample size (Lévy & Iglesias, 2006), however, the differences in CFI are irrelevant ($\Delta$$CFI < .01$ for universities and $\Delta$$CFI < .001$ for gender). Being less than 0.01, this allows us to assume configurational and metric invariance (Cheung & Rensvold, 2002) (see Table 4).

Establishment of Cut-Off Scores

Our cluster analysis of the initial sample ($n_i=510$) yielded four clusters, clearly defined according to both their total empathy scores and the individual dimension’s scores, and cut-off scores were set using the upper limit of the mean in each cluster. For instance, for the total scale score ($E$), we defined cut-off values of 88, 108, and 121, which resulted in four levels of empathy: low (20–88 points), medium low (89–108 points), medium-high (109–121 points), and high (122–140 points). Since the minimum and maximum scores are 20 and 140 points, respectively, we employed these theoretical values to set limits on the extreme values of total empathy, which empirically ranged from 66 points (minimum) to 139 points (maximum). This criterion was also applied to the scores for each dimension of empathy. The descriptive statistics of each level of empathy and its dimensions are shown in Table 5.

To verify that the levels set for empathy and its dimensions had been well delineated and used to define the categories considered to be sufficiently different, we performed an analysis of variance to compare the means of each level for total empathy and each of its dimensions. Results revealed an adequate effect size ($\eta^2$, partial eta squared, and $R^2$, coefficient of determination), with statistically significant differences as follows: Empathy ($F(3, 847) = 1373.78, p = .0001$, $\eta^2 = .862$, $R^2 = .889$), CC ($F(3, 847) = 627.3, p = .0001$, $\eta^2 = .850$, $R^2 = .889$), PT ($F(3, 847) = 813.16, p = .0001$, $\eta^2 = .890$, $R^2 = .932$), and WIPS ($F(3, 847) = 571.97, p = .0001$, $\eta^2 = .849$, $R^2 = .931$). To verify that the levels set for empathy and its dimensions are equally relevant, we performed a hierarchical cluster analysis that yielded four a priori clusters. We first classified all cases according to the cluster to which they originally belonged, and then according to the cut-off scores set. With this information, we generated 2 × 2 tables comparing the means of each level for total empathy and each of its dimensions. Results revealed an adequate effect size ($\eta^2$, partial eta squared, and $R^2$, coefficient of determination), with statistically significant differences as follows: Empathy ($F(3, 847) = 1373.78, p = .0001$, $\eta^2 = .862$, $R^2 = .889$), CC ($F(3, 847) = 627.3, p = .0001$, $\eta^2 = .850$, $R^2 = .889$), and WIPS ($F(3, 847) = 571.97, p = .0001$, $\eta^2 = .849$, $R^2 = .931$). To verify that the levels set for empathy and its dimensions are equally relevant, we performed a hierarchical cluster analysis that yielded four a priori clusters. We first classified all cases according to the cluster to which they originally belonged, and then according to the cut-off scores set. With this information, we generated 2 × 2 tables comparing the means of each level for total empathy and each of its dimensions. Results revealed an adequate effect size ($\eta^2$, partial eta squared, and $R^2$, coefficient of determination), with statistically significant differences as follows: Empathy ($F(3, 847) = 1373.78, p = .0001$, $\eta^2 = .862$, $R^2 = .889$), CC ($F(3, 847) = 627.3, p = .0001$, $\eta^2 = .850$, $R^2 = .889$), and WIPS ($F(3, 847) = 571.97, p = .0001$, $\eta^2 = .849$, $R^2 = .931$).
Table 2. Standardized Factor Loadings of the Original JSE Model, Global Sample, and Models Generated by University and Gender.

| Factor | Item | Global sample | \( R^2 \) | UBO    | UM    | UA    | Female | Male |
|--------|------|---------------|----------|--------|--------|--------|--------|------|
| PT     | P2   | 0.533         | .284     | 0.647  | 0.429  | 0.483  | 0.425  | 0.591|
|        | P4   | 0.483         | .234     | 0.647  | 0.448  | 0.678  | 0.333  | 0.591|
|        | P5   | 0.411         | .169     | 0.548  | **0.290*** | **0.197*** | **0.341*** | 0.443|
|        | P9   | 0.541         | .293     | 0.563  | 0.533  | 0.568  | 0.589  | 0.522|
|        | P10  | 0.576         | .332     | 0.502  | 0.679  | 0.588  | 0.587  | 0.574|
|        | P13  | 0.546         | .299     | 0.466  | 0.668  | 0.488  | 0.592  | 0.521|
|        | P15  | **0.236***     | .056     | **0.223*** | **0.221*** | 0.462  | **0.180*** | **0.278***|
|        | P16  | 0.580         | .336     | 0.519  | 0.722  | 0.588  | 0.563  | 0.583|
|        | P17  | **0.357***     | .127     | **0.286*** | 0.492  | **0.325*** | **0.379*** | **0.311***|
|        | P20  | 0.447         | .200     | **0.394*** | 0.586  | **0.366*** | 0.415  | 0.469|
| CC     | P1   | 0.581         | .338     | 0.543  | 0.491  | 0.404  | 0.636  | 0.542|
|        | P7   | 0.738         | .545     | 0.748  | 0.608  | 0.537  | 0.719  | 0.749|
|        | P8   | 0.730         | .533     | 0.758  | 0.637  | 0.531  | 0.730  | 0.737|
|        | P11  | 0.728         | .530     | 0.738  | 0.724  | 0.877  | 0.678  | 0.769|
|        | P12  | 0.699         | .489     | 0.749  | 0.664  | 0.497  | 0.691  | 0.712|
|        | P14  | 0.790         | .623     | 0.813  | 0.739  | 0.701  | 0.777  | 0.794|
|        | P18  | **0.330***     | .109     | 0.405  | **0.319*** | **0.189*** | **0.311*** | **0.368***|
|        | P19  | 0.485         | .235     | 0.443  | **0.272*** | **0.342*** | 0.497  | 0.423|
| WIPS   | P3   | 0.768         | .590     | 0.775  | 0.538  | 0.638  | 0.786  | 0.757|
|        | P6   | 0.671         | .450     | 0.530  | 0.774  | 0.753  | 0.721  | 0.608|

Note. \( R^2 \) = squared multiple correlations; UBO = Universidad Bernardo O’Higgins; UM = Universidad Mayor; UDA = Universidad de Atacama; CC = compassionate care; PT = perspective taking; WIPS = walking in patient’s shoes; E = empathy. **bold** indicates factor weights \( < 0.40 \).

Table 3. Jefferson Empathy Scale CFA Goodness-of-Fit Indices for Each University, Gender, and Total Sample.

| CFA model | \( \chi^2 \) | df | \( p \) | \( \chi^2/df \) | GFI | AGFI | SRMR | CFI | RMSEA [90% CI] |
|-----------|-------------|----|--------|----------------|-----|------|------|-----|----------------|
| UBO       | 242.676     | 162 | .000   | 1.498          | .913 | .887 | .0641 | .942 | .044 [.032; .056] |
| UM        | 211.162     | 161 | .005   | 1.312          | .836 | .786 | .0794 | .904 | .058 [.033; .078] |
| UDA       | 221.003     | 163 | .002   | 1.356          | .810 | .756 | .0839 | .853 | .064 [.040; .084] |
| Female    | 250.010     | 163 | .000   | 1.534          | .944 | .927 | .0468 | .955 | .036 [.027; .045] |
| Male      | 295.605     | 160 | .000   | 1.848          | .937 | .918 | .0504 | .943 | .044 [.036; .052] |
| Total     | 373.066     | 164 | .000   | 2.275          | .958 | .946 | .0428 | .951 | .039 [.034; .044] |

Note. UBO = Universidad Bernardo O’Higgins; UM = Universidad Mayor; UDA = Universidad de Atacama; GFI = Goodness of Fit Index; AGFI = adjusted goodness of fit index; SRMR = root mean square of standardized residuals; CFI = comparative fit index; RMSEA = root mean square error of approximation; CI = confidence interval.

Establishment of the Norm

Table 7 presents the percentiles of interest for each level of empathy and its dimensions. This information makes it possible to classify each participant as belonging to an empathy level and improves our interpretation of their score based on a normative sample.

Discussion

Regarding the students’ mean scores, data are not available for comparison with other physical therapy samples, but when compared with medical students from various other national and international studies, they present a mean score slightly lower than 112 points, with a standard deviation of around 12 (Hojat et al., 2018). There is a higher mean in women than in men, with a small effect size, \( d = 0.25 \) (Cohen, 1988, 1992), indicative that said difference from a practical or clinical perspective would not be highly relevant.

The reliability of the measurement (Cronbach’s \( \alpha = .835 \)) slightly exceeds other estimates made from a variety of university student samples in Chile and in other countries, where Cronbach’s alpha values have ranged from .70 to .80 with an average of .78 and an intraclass correlation of .835, indicative of good reliability (Hojat, 2018; Koo & Li, 2016). The JSE is a reliable measure for use with physical therapy students.

By studying the factor structure of the JSE, a three-factor model was obtained that fit the data well enough. The model confirmed all 10 items of the Perspective Taking factor, with
factor loadings equal to or greater than 0.24, and low factor loadings (< 0.40) on items 15 and 17, with a coefficient $\alpha = .73$. The Compassionate Care factor included eight items with factor loadings equal to or greater than 0.33, with low factor loading on item 18 and an $\alpha = .84$. The third factor, Walking in Patient’s Shoes, included two items with factor loadings of 0.67 and 0.77 and a coefficient $\alpha = .69$.

This three-factor model agrees with those reported for medical students in the US (Hojat et al., 2018), Spain (Ferreira-Valente et al., 2016), and Turkey (Bilgel & Ozcakir, 2017). The low internal consistency of the fourth and final factor is, we believe, explained by the small number of items that comprise it; ideally, a minimum of three items is required to stably determine a factor (Velicer & Fava, 1998). Three, or even better four, items per factor would significantly increase internal consistency (Ferrando & Anguiano-Carrasco, 2010). Despite the low factor loadings of three items (items 15 and 17 in PT, and 18 in CC), their presence does not damage the reliability of the measurement, although they have high error variance (with squared multiple correlations between .056 and .127).

Currently, the measurement of differences in levels of empathy is made using statistical estimates (Ye et al., 2020; Yuguero et al., 2019) that do not provide information about the qualitative changes observed when empathy is gained. The observed results of the factor model formed the basis for the determination of cut-off points. As a consequence, the establishment of these cut-off values provides a possible reference by which to establish comparisons between the empathy values observed not only in different schools of the same discipline within a country, but between countries and beyond this case of Chilean physical therapy students. Additionally, they serve as a reference point from which to measure the qualitative effect of a given empathic intervention; to discern whether interventions can effect change from one category to another or determine if such interventions have produced changes only within the same category.

Sensitivity exceeded 90% in two of the three cut-off points, satisfactory for an instrument, which measures a cognitive-affective psychological attribute, falling slightly only for the highest empathy categories (to 0.728) where it generated a false negative rate of 27.2%. We correctly identified those who possessed high empathy with a specificity that reached 98.7%. Without a doubt, further investigations into a criterion to permit the identification of those with high and low empathy are required.

Finally, the results of this study of Chilean physical therapy students provide evidence that allows inferring adequate psychometric properties of the JSE in this particular population, which is consistent with the evidence observed in samples of medical and dental students from Chile and Latin America. The three-factor, 20-item measurement model was reasonably fitted to the data, with satisfactory goodness-of-fit indicators, confirming the factor structure.

Thus, evidence was also obtained regarding factor invariance by gender, which indicated that the measure of empathy is equivalent in male and female students of physical therapy, favoring the comparison of the measurements by gender. The clusters generated provide cut-off points that assess empathy and its components categorically, and examine its changes, permitting potential comparisons between student populations, facilitating the interpretation of this variable, and ultimately simplifying decision-making processes.

A limitation of this study, due to a lack of representativeness in the sample used, is that its results cannot be generalized to all physical therapy students in Chile. To mitigate this, we selected samples from three different geographical areas of Chile.

This study includes elements, which contribute to the use of the JSE in professional contexts beyond the wide use it has been given in research. The psychometric properties of the instrument were reviewed, and norm and cut-off points were established which may be of wide and straightforward use for those health professionals who must measure empathy. It also provides sensitivity and specificity values, which permit

---

Table 4. Goodness of Fit of the Confirmatory Multigroup Factor Model According to University and Gender in Successive Nested Models.

| Model                              | $\chi^2$ | df  | p    | $\Delta\chi^2$ | $\Delta$df | p    | CFI  | $\Delta$CFI |
|-----------------------------------|----------|-----|------|----------------|------------|------|------|-------------|
| Invariance by University          |          |     |      |                |            |      |      |             |
| Base model/configural invariance  | 849.49   | 492 | .000 |                |            |      | .914 |             |
| Metric invariance                 | 914.611  | 526 | .000 | 65.121         | 34         | .001 | .906 | .008        |
| Structure invariance covariances  | 967.965  | 538 | .000 | 53.354         | 12         | .000 | .896 | .010        |
| Invariance by gender              |          |     |      |                |            |      |      |             |
| Base model/configural invariance  | 612.610  | 328 | .000 |                |            |      | .934 |             |
| Metric invariance                 | 635.530  | 345 | .000 | 22.921         | 17         | .152 | .932 | .002        |
| Structure invariance covariances  | 645.130  | 351 | .000 | 9.6            | 6          | .143 | .931 | .001        |

Note. $\Delta\chi^2$ difference between the $\chi^2$ values, $\Delta$df difference between degrees of freedom; CFI = comparative fit index; $\Delta$CFI = difference between the comparative fit index.
its use as a diagnostic test. The wide use of the JSE in the context of research constitutes a foundation for its future development and, in this context, it appears to us that the continued establishment of test norms should proceed, and national norms for each country in which the test is used should be found which consider different health professions and health science training areas. From the psychometric perspective, the performance of some items require further investigation: in particular those in which several studies have shown low factorial loads (items 15, 17, and 18), as well as an evaluation of the real contribution the specification of the respective factors makes. Equally, it appears reasonable to pay attention to the WIPS factor’s significance in the construction of empathy; a factor, which in various studies has shown low relative reliability with respect to the other two of the scale’s factors or, indeed, has damaged the general goodness-of-fit of the original three-base-factor model, potentially by testing new items, which may improve the measurement of the factor, looking to widen it to three or four elements.

**Conclusion**

Despite the wide use of the JSE at a global level, the establishment of norms has not advanced sufficiently to allow the interpretation of a person’s scores in relation to a representative sample of their population. This article proposes a norm for Chilean physical therapy students, and was able to position them relative to others by translating their test grade into a percentage value. This use of the raw empathy score to place them in the categories of high, medium-high, medium-low, and low with respect to the total empathy score or its dimensions, and contribution of cut-off points of adequate sensitivity and specificity widens the possibilities of use of the JSE. Confirming a factorial structure that contributes to the validity of the construct, along with adequate internal validity.

### Table 5. Descriptive Statistics of Empathy Levels and Its Components.

| Level | n   | Min. | Max.  | M    | SD   | Median | Huber M-estimator |
|-------|-----|------|-------|------|------|--------|-------------------|
| E     |     |      |       |      |      |        |                   |
| High  | 125 | 122  | 139   | 127.52 | 3.75 | 127    | 126.99           |
| Medium-high | 148 | 109  | 121   | 115.64 | 3.56 | 116    | 115.91           |
| Medium-low | 138 | 89   | 108   | 100.06 | 5.48 | 101    | 100.66           |
| Low   | 98  | 66   | 88    | 81.69  | 4.60 | 82.5   | 82.29            |
| PT    |     |      |       |      |      |        |                   |
| High  | 323 | 60   | 70    | 64.97  | 3.07 | 65     | 66.96            |
| Medium-high | 90  | 55   | 59    | 56.99  | 1.30 | 57     | 62.21            |
| Medium-low | 78  | 45   | 54    | 50.74  | 2.48 | 51     | 57.89            |
| Low   | 18  | 38   | 44    | 42.11  | 1.88 | 42.5   | 54.90            |
| CC    |     |      |       |      |      |        |                   |
| High  | 319 | 38   | 56    | 46.93  | 4.39 | 47     | 47.12            |
| Medium-high | 107 | 26   | 37    | 32.40  | 3.01 | 33     | 32.75            |
| Medium-low | 45  | 19   | 25    | 22.27  | 2.02 | 23     | 22.35            |
| Low   | 38  | 8    | 18    | 14.92  | 2.33 | 15     | 15.10            |
| WIPS  |     |      |       |      |      |        |                   |
| High  | 55  | 13   | 14    | 13.60  | 0.49 | 14     | *                |
| Medium-high | 215 | 8    | 12    | 9.53   | 1.43 | 9      | 9.29             |
| Medium-low | 210 | 4    | 7     | 5.51   | 1.08 | 6      | 5.55             |
| Low   | 29  | 2    | 3     | 2.59   | 0.50 | 3      | *                |

Note. E = empathy, PT = perspective taking, CC = compassionate care, WIPS = walking in patient’s shoes, M = mean, SD = standard deviation. *Some M estimators cannot be calculated due to the highly centralized distribution around the median.

### Table 6. Sensitivity, Specificity, and Characteristics of the Full Scale.

| Cut-off points in the HPS | Se   | Sp   | FN  | FP  | OR  | CI 95% (OR) | PP [CI 95%] |
|---------------------------|------|------|-----|-----|-----|-------------|-------------|
| 88                        | 0.951| 0.810| 4.9%| 19% | 5   | 3.33–7.53   | 75% [67%–82%]|
| 108                       | 0.930| 0.522| 7%  | 47.8%| 1.95| 1.64–2.31   | 55% [51%–60%]|
| 121                       | 0.728| 0.987| 27.2%| 1.3%| 56 | 7.97–394   | 99% [91%–100%]|

Note. Se = sensitivity; Sp = specificity; FN = false negative; FP = false positive; OR = odds ratio (positive likelihood ratio); CI = confidence interval; PP = posterior probability.
### Table 7. Results of the Estimation of the Percentiles of the Empathy and Its Components in Each Observed Cluster and Their Minimum and Maximum Values (Cut-off Points).

| Rank | Level   | 5  | 10 | 25 | 50 | 75 | 90 | 95 |
|------|---------|----|----|----|----|----|----|----|
|      |         |    |    |    |    |    |    |    |
| Empathy |         |    |    |    |    |    |    |    |
| 122–140 | High   | 122 | 123 | 125 | 127 | 130 | 133 | 134 |
| 109–121 | Medium-high | 110 | 110 | 113 | 116 | 119 | 120 | 120 |
| 89–108 | Medium-low | 91  | 92  | 96  | 101 | 105 | 107 | 108 |
| 20–88 | Low     | 73  | 75  | 79  | 83  | 86  | 87  | 88  |
| Compassionate care |         |    |    |    |    |    |    |    |
| 38–56 | High   | 39  | 40  | 44  | 47  | 50  | 52  | 53  |
| 26–37 | Medium-high | 27  | 27  | 30  | 33  | 35  | 36  | 37  |
| 19–25 | Medium-low | 19  | 19  | 21  | 23  | 24  | 25  | 25  |
| 8–18 | Low     | 11  | 11  | 14  | 15  | 16  | 18  | 18  |
| Perspective taking |         |    |    |    |    |    |    |    |
| 60–70 | High   | 62  | 63  | 65  | 67  | 69  | 70  | 70  |
| 55–59 | Medium-high | 54  | 56  | 59  | 62  | 66  | 68  | 69  |
| 45–54 | Medium-low | 48  | 50  | 53  | 58  | 62  | 66  | 67  |
| 10–44 | Low     | 42  | 43  | 48  | 55  | 61  | 64  | 65  |
| Walking in patient’s shoes |         |    |    |    |    |    |    |    |
| 13–14 | High   | 13  | 13  | 13  | 14  | 14  | 14  | 14  |
| 8–12 | Medium-high | 8   | 8   | 8   | 9   | 11  | 12  | 12  |
| 4–7 | Medium-low | 4   | 4   | 5   | 6   | 6   | 7   | 7   |
| 2–3 | Low     | 2   | 2   | 2   | 2   | 3   | 3   | 3   |

consistency, indicates the reliability of the measurement for use with physical therapy students.

### 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
Victor Díaz-Narváez [https://orcid.org/0000-0002-5486-0415](https://orcid.org/0000-0002-5486-0415)

### References
Avila, V., Carrasco, C., Osorio, L., Calzadilla-Núñez, A., & Díaz-Narváez, V. P. (2020). Estudio transversal de empatía con el paciente en estudiantes de kinesiología (Cross-sectional study of empathy with the patient in physical therapy students). *Educación Médica Superior*, 34(2), e1919.

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88(3), 588–606.

Bilgel, N., & Ozçakir, A. (2017). Turkish version of the Jefferson scale of empathy psychometric properties. *European Scientific Journal*, 13(20), 101–111. https://doi.org/10.19044/esj.2017.v13n20p101

Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research*, 21, 230–258.

Cajal, B., Gervilla, E., & Palmer, A. (2012). When the mean fails, use an M-estimator. *Anales de Psicología*, 28(1), 281–288.

Calzadilla-Núñez, A., Díaz-Narváez, V., Dávila-Pontón, Y., Aguilera-Muñoz, J., Fortich-Mesa, N., Aparicio-Mareno, D., & Reyes-Reyes, A. (2017). Empathic erosion during medical training according to gender: Cross-sectional study. *Archivos Argentinos de Pediatría*, 115(6), 556–561. https://doi.org/10.5546/aap.2017.eng.556

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling A Multidisciplinary Journal*, 9(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155–159.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334. https://doi.org/10.1007/BF02310555

Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, 1(1), 16–29. https://doi.org/10.1037/1082-989x.1.1.16

Díaz-Narváez, V. P., Alonso-Palacio, L. M., Caro, S., Silva, M., Arboleda-Castillo, J., Bilbao, J. L., et al. (2017). Compassionate care component of the construct empathy in medical students in Colombia and Dominican Republic. *Acta Médica Mediterránea*, 33(1), 115–121. https://doi.org/10.19193/0393-6384.2017.1_018

Díaz-Narváez, V. P., & Calzadilla-Núñez, A. (2019). Ecualización de la empatía en estudiantes de dos sedes diferentes en una facultad de odontología de una universidad chilena (Empathy equalization in students from two different campuses in a Chilean Faculty of Dentistry). *Revista Médica. Rosario*, 85, 20–26.

Díaz-Narváez, V. P., Salas-Alarcón, D., Bracho-Milic, F., & Ocaranza-Ozímica, J. (2015). Empatía en estudiantes de fisioterapia. Universidad Mayor, sede Temuco, Chile (Empathy in students of specialty in physical therapy, Universidad Mayor in Temuco, Chile). *Revista Ciencias de la Salud*, 12(3), 383–393. https://doi.org/10.12804/revsalud13.03.2015.05

Ferrando, P. J., & Anguiano-Carrasco, C. (2010). El análisis factorial como técnica de investigación en psicología (Factor analysis as a technique in psychological research). *Papeles del Psicólogo*, 31(1), 18–33.

Ferreira-Valente, A., Costa, P., Eldorduy, M., Virumbrales, M., & Costa, M. J. (2016). 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 Medical Education*, 16, 242. https://doi.org/10.1186/s12909-016-0763-5

Galán, J. M. G., Serrano, R. R., Martin, M. S. M., & Fernández, J. M. A. (2014). Descenso de empatía en estudiantes de enfermería y análisis de posibles factores implicados. (Decreasing empathy in nursing students and analysis of possible factors involved). *Psicologíca Educativa*, 20(1), 53–60. https://doi.org/10.1016/j.pse.2014.05.007

González-Martínez, F., Tirado-Amador, L., Bueno-Hernández, J., Chica-Duque, B., & Díaz-Narváez, V. (2018). Changes in empathy levels on dentistry’s students of Public University...
in Cartagena City, Colombia. *Pesquisa Brasileira em Odontopediatria e Clínica Integrada*, 18(1), 1–12. https://doi.org/10.4034/PBOCI.2018.181.44

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2013). *Multivariate data analysis* (7th ed.). Pearson Education.

Hojat, M., Gonnella, J. S., Nasca, T. J., Mangione, S., Veloski, J. J., & Magee, M. (2002). The Jefferson scale of physician empathy: Development and preliminary psychometric data. *Academic Medicine*, 77(10 Suppl.), S58–S60.

Hojat, M., De Santis, 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. https://doi.org/10.1007/s10459-018-9839-9

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., Mangione, S., Nasca, T. J., Cohen, M. J. M., Gonnella, J. S., Erdmann, J. B., Veloski, J. J., & Magee, M. (2001). The Jefferson scale of physician empathy: Development and preliminary psychometric data. *Educational and Psychological Measurement*, 61, 349–365.

Hojat, M., Shannon, S. C., DeSantis, J., Speicher, M. R., Bragan, L., & Calabrese, L. H. (2019). Empathy in medicine national Norms for the Jefferson scale of empathy: A nationwide project in osteopathic medical education and empathy (POMEE). *Journal of Osteopathic Medicine*, 119(8), 520–532. https://doi.org/10.7556/joa0.2019.091

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), 1–55. https://doi.org/10.1080/10705519909540118

Hunt, P. A., Denieffe, S., & Gooney, M. (2017). Burnout and its relationship to empathy in nursing: A review of the literature. *Journal of Research in Nursing*, 22(1-2), 7–22. https://doi.org/10.1177/174498116678902

Jöreskog, K. G. (1971). Simultaneous factor analysis in several populations. *Psychometrika*, 36, 409–426. https://doi.org/10.1007/BF02291366

Kline, R. B. (2005). *Principles and practice of structural equation modeling*. The Guilford Press.

Kolmogorov, A. (1933). Sulla determinazione empirica di una legge di distribuzione. *Giornale dell’ Istituto Italiano Degli Attuari*, 4, 83–91.

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15, 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Lee, Y., & Seomun, G. (2016). Development and validation of an instrument to measure nurses’ compassion competence. *Applied Nursing Research*, 30, 76–82. https://doi.org/10.1016/j.apmr.2015.09.007

Levene, H. (1960). Robust tests for the equality of variances. In En J. Olkin (Ed.), *Contributions to probability and statistics* (pp. 278–292). Stanford University Press.

Lévy, J. P., & Iglesias, V. (2006). Invarianza causal con muestras múltiples. In F. Losada Pérez, J. Varela Mallou, & J.-P. Lévy Mangin (Eds.), *Modelización con Estructuras de Covarianzas en Ciencias Sociales* (pp. 279–318). Nethibio.

Madera-Anaya, M., Tirado-Amador, L., & González-Martínez, F. (2016). Factores relacionados con la empatía en estudiantes de Enfermería de la Universidad de Cartagena. (Factors related to empathy in medical students at the University of Cartagena). *Enfermería Clínica*, 26(5), 282–289.

McDonald, R. P. (1999). *Test theory: A unified treatment*. Lawrence Erlbaum.

Mohamed, M., & Shoukri. (2004). *Measures of interobserver agreement*. Chapman & Hall/CRC.

Pastén-Hidalgo, W. F., van Niekerk-Bakit, N. A., Calzadilla-Núñez, A., Aguilera-Olivares, F., & Díaz-Narváez, V. P. (2019). Empatía en estudiantes de fisioterapia: Tendencia por curso y género. Declinación empática. (Empathy in physiotherapy students: Trend by course and gender, Empathy decline). *Fisioterapia*, 41(5), 250–257. https://doi.org/10.1016/j.ft.2019.05.004

Preusche, L., & Lamm, C. (2016). Reflections on empathy in medical education: What can we learn from social neurosciences? *Advances in Health Sciences Education*, 21(1), 235–249. https://doi.org/10.1007/s10459-015-9581-5

Rueckert, L., Branch, B., & Doan, T. (2011). Are gender differences in empathy due to differences in emotional reactivity? *Psychology*, 2(6), 574–578. https://doi.org/10.4236/psych.2011.26088

Stevens, J. P. (1992). *Applied multivariate statistics for the social sciences*. Erlbaum.

Svenaeus, F. (2016). The phenomenology of empathy: A Steinian emotional account. *Phenomenology and the Cognitive Sciences*, 15(2), 227–245. https://doi.org/10.1007/s11097-014-9411-x

Velicer, W. F., & Fava, J. L. (1998). Affects of variable and subject sampling on factor pattern recovery. *Psychological Methods*, 3(2), 231–251. https://doi.org/10.1037/1082-989X.3.2.231

Ye, X., Guo, H., Xu, Z., & Xiao, H. (2020). Empathy variation among osteopathic medicine students: Trend by course and gender, Empathy decline). *Fisioterapia*, 41(5), 250–257. https://doi.org/10.1016/j.ft.2019.05.004

Ye, X., Guo, H., Xu, Z., & Xiao, H. (2020). Empathy variation among osteopathic medicine students. *Journal of Osteopathic Medicine*, 119(8), S58–S60.

Zúñiga, Y., De Dios, B., Craftos, P., Sánchez, R., & Alcalde, A. (2018). El estudio del apoyo empático y de las variables relacionadas con el bienestar emocional en estudiantes de Fisioterapia (Comparación de género y nivel de estudios). *Fisioterapia*, 30(2), 227–245. https://doi.org/10.1016/j.fisio.2018.09.007

Ye, X., Guo, H., Xu, Z., & Xiao, H. (2020). Empathy variation among osteopathic medicine students. *Journal of Osteopathic Medicine*, 119(8), S58–S60.