Jefferson.

Thomas Jefferson University

Jefferson Digital Commons

Department of Psychiatry and Human Behavior Faculty Papers

Department of Psychiatry and Human Behavior

Eleven Years of Data on the Jefferson Scale of Empathy-Medical Student Version (JSE-S): Proxy Norm Data and Tentative Cutoff Scores.

Mohammadreza Hojat

Joseph S. Gonnella

Follow this and additional works at: https://jdc.jefferson.edu/phbfp

Part of the Medical Education Commons

Let us know how access to this document benefits you

Medical Principles and Practice

Eleven Years of Data on the Jefferson Scale of Empathy-Medical Student Version (JSE-S): Proxy Norm Data and Tentative Cutoff Scores

Mohammadreza Hojat   Joseph S. Gonnella

Center for Research in Medical Education and Health Care, Sidney Kimmel (formerly Jefferson) Medical College at Thomas Jefferson University, Philadelphia, Pa., USA

Abstract

Objective: This study was designed to provide typical descriptive statistics, score distributions and percentile ranks of the Jefferson Scale of Empathy-Medical Student version (JSE-S) of male and female medical school matriculants to serve as proxy norm data and tentative cutoff scores. ***Subjects and Methods:*** The participants were 2,637 students (1,336 women and 1,301 men) who matriculated at Sidney Kimmel (formerly Jefferson) Medical College between 2002 and 2012, and completed the JSE at the beginning of medical school. Information extracted from descriptive statistics, score distributions and percentile ranks for male and female matriculants were used to develop proxy norm data and tentative cutoff scores. ***Results:*** The score distributions of the JSE tended to be moderately skewed and platykurtic. Women obtained a significantly higher mean score (116.2 ± 9.7) than men (112.3 ± 10.8) on the JSE-S ($t_{2,635} = 9.9$, $p < 0.01$). It was suggested that percentile ranks can be used as proxy norm data. The tentative cutoff score to identify low scorers was ≤95 for men and ≤100 for women. ***Conclusions:*** Our findings provide norm data and cutoff scores for admission decisions under certain conditions and for identifying students in need of enhancing their empathy.
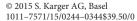
© 2015 S. Karger AG, Basel

Introduction

Empathy is an elusive concept. There are different descriptions or definitions of empathy in social psychology, but a more focused and relevant definition is needed in the context of education and care of patients in the health professions. Empathy is defined in the context of health professional education and patient care as: 'predominantly a *cognitive* (as opposed to affective or emotional) attribute that involves *understanding* (as opposed to feeling) of the patient's pain, experiences, concerns, and perspectives combined with a capacity to *communicate* this understanding and an *intention to help*' [1, 2].

Prior to the development of the Jefferson Scale of Empathy (JSE), no psychometrically sound instrument was available to specifically measure empathy in patient care. Although a few research tools existed for measuring empathy in the general population [1], none of these was content-specific and context-relevant to patient

Dr Mohammadreza Hojat
Sidney Kimmel (formerly Jefferson) Medical College
1015 Walnut Street, Curtis Building, 3rd Floor, Suite 320B
Philadelphia, PA 19107 (USA)
E-Mail mohammadreza.hojat@jefferson.edu

care. More than a decade ago, we recognized a need for a psychometrically sound instrument to measure empathy in the context of health professional education and patient care. In response to that need, we developed the JSE [1–4]. There are 3 versions of the JSE: (1) the HP-version: for administration to physicians and practitioners of all health professions, (2) the S-version: for administration to medical students and (3) the HPS-version: for administration to students in all health professions other than medicine. The three versions are very similar in content with only slight modifications in wording to make the text more appropriate for the target population.

Evidence in support of the psychometric properties of the JSE has been reported [3–6]. The JSE has been widely used for different health professional students and practitioners in the USA and abroad, has been translated into 47 languages and is used in more than 70 countries [7]. There is a large volume of research by national and international researchers who have used the JSE with some consistent findings including gender difference in favor of women [3–5]. Furthermore, in most of these studies, it has been noticed that high JSE scorers were more likely than low scorers to pursue the so-called 'people-oriented' specialties (e.g. general internal medicine, family medicine and pediatrics) as opposed to 'technology- or procedure-oriented' specialties (e.g. pathology, radiology, anesthesiology and surgery) [3–5].

Study Objective

As the developers of the JSE, we have been asked frequently by potential users about the availability of norm data and cutoff scores for identifying high and low scorers. For the development of norm tables and determining cutoff scores, large and representative samples from the target populations are needed. However, as an initial step, it seemed reasonable to use large samples from a typical medical school to provide proxy norm data and tentative cutoff scores for the JSE S-version (JSE-S). We designed this study in response to a need for norm data and cutoff scores.

Subjects and Methods

Participants included 2,637 students (1,336 women and 1,301 men) who matriculated at Sidney Kimmel (formerly Jefferson) Medical College at Thomas Jefferson University in Philadelphia, Pa., USA, between 2002 and 2012, and who had completed the JSE-S at the beginning of medical school (representing a 94% response rate).

Table 1. Frequency and percentage distributions of the study sample by matriculation year and gender

Matriculation year	Men, n (%)	Women, n (%)	Total
2002	120 (54)	101 (46)	221 (100)
2003	105 (48)	113 (52)	218 (100)
2004	103 (46)	121 (54)	224 (100)
2005	126 (51)	121 (49)	247 (100)
2006	107 (43)	140 (57)	247 (100)
2007	132 (53)	116 (47)	248 (100)
2008	120 (51)	117 (49)	237 (100)
2009	111 (46)	128 (54)	239 (100)
2010	124 (49)	128 (51)	252 (100)
2011	125 (50)	127 (50)	252 (100)
2012	128 (51)	124 (49)	252 (100)
Total	1,301 (49)	1,336 (51)	2,637 (100)

$\chi^2_{10} = 9.8$ (p = 0.45, nonsignificant).

Being part of the Jefferson Longitudinal Study of Medical Education Outcomes, the study had been approved by Thomas Jefferson University's Institutional Review Board, and no consent form was required. The hard copy of the JSE-S was administered to the incoming medical students each year at the orientation day for the entering classes of 2002–2006, and it was administered online for the entering classes of 2007–2012. For examining the validity of the cutoff scores, we used average clinical competence ratings in 6 third-year core clerkships (family medicine, internal medicine, obstetrics/gynecology, pediatrics, psychiatry and surgery) and average ratings given by postgraduate training program directors at the completion of the first postgraduate year of postgraduates' clinical competence for the factors the 'art' and 'science' of medicine [8]. Participation was voluntary.

Statistical Analysis

For the purpose of determining the tentative cutoff scores to identify high and low scorers, we arbitrarily chose 2 points on the score distributions. To identify high scorers, we chose a point on the score distribution which was one and half standard deviation above the mean score. To identify low scorers, we chose another point which was one and half standard deviation below the mean score. Due to gender differences on the JSE [3–5], the cutoff scores for men and women were calculated separately from their respective score distributions.

We compared performance measures and the clinical competence ratings among high, moderate and low JSE scorers to examine the validity of the cutoff scores. In addition to descriptive statistics, we used the χ^2 test, the one-tailed Student t test and analysis of variance (ANOVA) to examine associations between the JSE scores and the criterion measures. We also calculated Cohen's *d* as an estimate of the effect size [9, 10]. The effect size values <0.25 were considered negligible, around 0.50 as moderate and >0.75 as large [9, 10]. We used SAS version 9.2 for Windows (SAS, Cary, N.C., USA) for the statistical analyses.

Table 2. The mean, SD, range, skewness and kurtosis indices, and reliability (Cronbach α) coefficients of the JSE by matriculating classes and summary results of statistical analysis

Matriculating Class	Students, n	Mean ± SD	Median	Range	Skewness	Kurtosis	Cronbach's α
2002	221	114.1±9.9	114	81–137	–0.24	0.04	0.80
2003	218	113.9±10.0	115	75–140	–0.44	0.52	0.79
2004	224	115.9±9.8	117	82–140	–0.35	0.12	0.78
2005	247	114.5±9.7	116	82–133	–0.66	0.46	0.78
2006	247	114.8±9.4	115	86–135	–0.46	0.19	0.75
2007	248	114.6±10.6	114	71–136	–0.47	0.74	0.81
2008	237	113.5±12.1	114	52–140	–0.92	2.66	0.84
2009	239	113.2±11.3	113	73–140	–0.28	0.05	0.84
2010	252	113.8±10.7	114	70–140	–0.62	0.88	0.81
2011	252	114.1±10.1	116	76–140	–0.57	0.79	0.79
2012	252	114.8±10.6	116	79–140	–0.65	0.90	0.81
Total	2,637	114.3±10.4	115	52–140	–0.56	0.92	0.80

$F_{10, 2,626} = 1.2$ (p = 0.29, nonsignificant).

Results

For the entire sample, the mean score for age was 23.4 ± 2.4 years (mean 23.5 ± 2.4 for men and 23.2 ± 2.3 for women). The gender composition of the sample by matriculation year is presented in table 1. The number of women varied from 101 (46%) in 2002 to 140 (57%) in 2006. The corresponding figures for men were 120 (54%) and 107 (43%), respectively. The results of the χ^2 test showed no significant difference in gender composition in different matriculation years ($\chi^2_{10} = 9.8$, p = 0.45).

Descriptive Statistics

The mean, SD, median, score range, skewness and kurtosis indices of the JSE-S for the entire sample and for the matriculants of each year are presented in table 2. The JSE mean score for the entire sample was 114.3 ± 10.4, which varied from a low of 113.2 ± 11.3 for the matriculants of 2009 to a high of 115.9 ± 9.8 for the matriculants of 2004. The ANOVA used to test the significance of JSE-S mean scores of matriculants from different years did not reveal any statistically significant differences ($F_{10, 2,626} = 1.2$, p = 0.29).

The skewness index was negative for the entire sample (–0.56) and for each matriculating year [range: –0.92 (for matriculants of 2008) to –0.24 (for matriculants of 2002) and median = –0.53]. The kurtosis for the entire sample was 0.92 [range: 0.04 (for matriculants of 2002) to 2.66 (for matriculants of 2008) and median = 0.52] (table 2).

Table 3. Gender differences on the JSE scores by matriculating classes

Class	Men		Women		t	Effect size[a]
	n	mean ± SD	n	mean ± SD		
2002	120	112.3±10.1	101	116.3±9.3	3.1**	0.41
2003	105	111.5±10.8	113	116.1±8.6	3.4**	0.46
2004	103	113.7±9.6	121	117.7±9.6	3.1**	0.43
2005	126	112.1±10.2	121	117.0±8.6	4.1**	0.52
2006	107	112.8±9.2	140	116.3±9.3	3.0**	0.37
2007	132	112.8±11.7	116	116.6±8.6	2.7**	0.40
2008	120	112.2±11.9	117	114.8±12.3	1.6†	0.21
2009	111	109.8±11.5	128	116.1±10.3	4.5**	0.57
2010	124	111.7±10.8	128	115.8±10.4	3.1**	0.38
2011	125	112.6±11.0	127	115.6±9.0	2.4*	0.30
2012	128	113.4±10.9	124	116.4±10.2	2.3*	0.28
Total	1,301	112.3±10.8	1,336	116.2±9.7	9.9*	0.40

* p < 0.05, ** p < 0.01, † p = 0.10. [a] Cohen's *d*.

Internal Consistency Reliability

Cronbach's α coefficient for the entire sample was 0.80 [range: 0.75 (for matriculants of 2006) to 0.84 (for matriculants of 2008 and 2009) and median = 0.80] (table 2).

Gender Difference on the JSE Scores

The gender differences on the JSE-S mean scores for men and women and for the entire sample as well as for

Hojat/Gonnella

Table 4. Frequency and percentage distributions and descriptive statistics of scores on the JSE-S by gender

Score interval	Men (n = 1,301)			Women (n = 1,336)			Total (n = 2,637)		
	frequency	cumulative frequency	percentile ranks	frequency	cumulative frequency	percentile ranks	frequency	cumulative frequency	percentile ranks
≤80	11	11	1%	5	5	<1%	16	16	<1%
81–85	8	19	1%	2	7	<1%	10	26	1%
86–90	22	41	2–3%	1	8	1%	23	49	2%
91–95	48	89	4–7%	21	29	2%	69	118	3–4%
96–100	87	176	8–13%	56	85	3–6%	143	261	5–10%
101–105	136	312	14–24%	89	174	7–13%	225	486	11–18%
106–110	214	526	25–40%	165	339	14–25%	379	865	19–33%
111–115	252	778	41–60%	258	597	26–45%	510	1,375	34–52%
116–120	232	1,010	61–78%	279	876	46–65%	511	1,886	53–71%
121–125	159	1,169	79–90%	221	1,097	66–82%	380	2,266	72–86%
126–130	91	1,260	91–97%	171	1,268	83–95%	262	2,528	87–96%
131–135	34	1,294	98–99%	56	1,324	96–99%	90	2,618	97–99%
>135	7	1,301	100%	12	1,336	100%	19	2,637	100%
Mean score[a]	112.3±10.8			116.2±9.7			114.3±10.4		
Median score	113			117			115		
SD	10.8			9.7			10.4		
Possible range	20–140			20–140			20–140		
Actual range	70–140			52–140			52–140		

[a] $t_{2,635} = 9.9$ (p < 0.0001 for testing the null hypothesis that JSE mean scores for men and women are not different).

each matriculating class are summarized in table 3. Women consistently and significantly (p < 0.01) obtained higher JSE-S mean scores than men, with the exception of the matriculating class of 2008 in which women's higher mean scores were not significantly different from those of men at the conventional level of statistical significance, i.e. usually p < 0.05 ($t_{235} = 1.6$, p = 0.07). The effect size estimates of the differences varied for different matriculating classes [range: 0.21 (for matriculants of 2008) to 0.57 (for matriculants of 2009)]. For the entire sample, the effect size estimate of gender difference was 0.40 ($t_{2,635} = 9.90$, p < 0.01).

Score Distributions and Percentile Ranks
Frequency distributions of the JSE scores and the percentile ranks for men, women and the entire sample are presented in table 4. The mean, median and SD for the entire sample were 114.3 ± 10.4, 115 and 10.4, respectively.

Tentative Cutoff Scores
The low and high cutoff scores for men were ≤95 and ≥127, respectively; the corresponding scores for women

were ≤100 and ≥129. These cutoff scores include approximately 7% of the top scorers and 7% of the bottom scorers in both the male and female samples. Results of ANOVA showed that differences on the clinical competence ratings and ratings of clinical competence in the 6 core clerkships were marginally significant ($F_{2, 2,284} = 2.57$, p = 0.07). In additional analyses, no statistically significant associations were found (p > 0.05) between low empathy scorers and the performance on objective licensing examinations of medical knowledge such as Step 1 (taken at the completion of the second medical school year) and Step 2 of the US Medical Licensing Examinations (taken in the fourth year of medical school).

Discussion

Our findings showed that the score distributions of the JSE were generally negatively skewed. Skewness index is a measure of symmetry in score distribution [11]. In a perfectly normal distribution, the skewness is close to zero. Negative skewness indicates that the peak of JSE-S score distributions tended to be to the right side of the

distribution (bulk of data to the side of higher scores); however, the magnitudes of the skewness indices suggest that distributions were just moderately skewed (distributions with skewness indices outside of the –1 to +1 range are considered highly skewed). Thus, our findings suggest that the JSE-S score distributions, although slightly skewed, do not substantially deviate from normal distributions.

Findings showed that the JSE score distributions tend to be platykurtic. Kurtosis is an index of the peak of score distribution [11]. Higher values indicate a higher peak and lower values indicate a flatter peak. Normal distributions have a kurtosis index close to 3 (mesokurtic), those >3 are high-peaked distributions (leptokurtic) and those with kurtosis <3 are flatter-peaked (platykurtic).

The examination of association between the categories of cutoff scores (i.e. high, moderate and low scorers) and performance measures revealed a consistent pattern of findings in the expected direction, in which the low scorers, when compared to the moderate and high scorers, received lower average ratings of clinical competence in the 6 third-year core clerkships as well as the ratings by the postgraduate program directors of the factors the 'art' and the 'science' of medicine [8]. However, the differences were only marginally significant (p < 0.07); one reason for this could be the exclusion of dropout students from the statistical analysis, which would lead to a narrower range of ratings that does not allow for the capture of the full range of relationships.

The frequency distributions, descriptive statistics and percentile ranks can serve as proxy norm data for matriculating students in any US medical school under the condition that the descriptive statistics and score distributions of the JSE for those schools are not substantially different from the data reported in table 4. For example, a score of 120 on the JSE-S obtained by a male matriculant would place him in the 78th percentile, and the same score obtained by a female matriculant would place her in the 65th percentile of the score distributions.

It is also interesting to note that we found no significant difference in JSE-S scores when comparing two types of test administrations, i.e. hard-copy testing for the classes of 2002–2006 and online testing for the classes of 2007–2012. This finding suggests that the type of test administration does not have any effect on the JSE-S scores. In addition, the stability of the empathy scores in different matriculating classes over the 11-year period of the study may suggest that the erosion of empathy in medical schools, as observed in other studies [2, 12, 13], is more likely due to the nature of educational programs, the

learning environment, a lack of positive role models and students' negative experiences in medical school [2] rather than to the methods of student selection or the types of student applying to medical schools at different periods of time.

Research findings where the JSE has been used have shown that empathy tends to decline as students make their way through medical school and schools for other health professionals [2, 12–15], and that empathy can be enhanced through educational programs that target medical students [16] as well as other students in the health professions [17]. In addition, research previously showed that empathy can be sustained in patient-care settings [18–20]. Given these research results, our findings regarding norm data and cutoff scores can be helpful for assessing the empathy of physicians-in-training for remedial programs.

More importantly, our previous findings, i.e. that physician empathy can positively predict optimal clinical outcomes in the control of diabetes (determined by the results of tests for hemoglobin A1c and low-density lipoprotein cholesterol) [21] and outcomes in diabetic patients (determined by hospitalization rates due to metabolic complications like a hyperosmolar state, diabetic ketoacidosis and coma) [22], suggest that empathy should also be considered as an essential component of overall physician competence. The findings strengthen the belief that the empathy of health-care providers is significantly linked to the outcomes of patients. Therefore, any approach to identify those who are in need of training to enhance empathy is beneficial to patient care.

Although the pattern of findings regarding empathy cutoff scores and assessment of clinical competence was in the direction expected, the associations did not reach the conventional levels of statistical significance (often p < 0.05). However, it is encouraging to observe a pattern of associations in the direction expected, given the time interval between administering the JSE (at the very beginning of medical school) and the assessment of clinical competence in medical school (3 years into medical school) and in postgraduate medical education (5 years after administration of the JSE).

Additional longitudinal cohort research is needed to further examine the associations between the suggested cutoff scores and assessment of clinical competence of medical students, in order to confirm the predictive validity of the cutoff scores. This is just the first step of a long journey for developing national and international norm tables. Indeed, further research is needed on representative samples of medical school matriculants in a

variety of medical schools and in different countries, so as to develop national and international norm tables and cutoff scores on the JSE for use by applicants to medical education academic centers for undergraduates and graduates.

Strengths and Limitations

The strengths of the study include its relatively large sample of 11 classes and the separate analyses conducted for men and women. A limitation is its single-institution feature, which may jeopardize the external validity or generalization of the findings. However, this limitation can be mitigated by the positive aspects of the study, i.e. its large sample and the 11-year span of data collection, as well as by the fact that Jefferson Medical College is typical of most 4-year allopathic medical schools in the USA with regard to the demographic composition of students and the specialty choices of graduates.

Implications

A potential implication of our findings is that the score distributions and percentile ranks reported here can be used as proxy norms for the purpose of comparing individual scores and determining the relative rank for male and female medical school matriculants (assuming that the score distributions and descriptive statistics of the medical school from which the JSE score is being compared are not substantially different from those reported in table 4). For example, the JSE-S score of a male matriculant in medical school 'X' that falls between 131 and 135 would place him in the top 98–99th percentile, and a score of a female matriculant from the same school that falls between 126 and 130 would place her in the 83–95th percentile (assuming similarity in the descriptive statistics and score distributions of the JSE in medical school 'X' with those reported in table 4).

The tentative cutoff scores suggested in this study are not absolutely definitive. We need data on well-validated criterion measures to examine the predictive validity of the cutoff scores. We also need more data from representative samples of medical schools at the national level to be able to develop national norm tables for male and female medical school matriculants. Using a similar approach, national (and international) norm tables could also be developed for students in other schools for health professionals (and in other countries) and for male and female doctors in different specialties. These ideas set an agenda for future research.

Conclusions

Our findings provide empirical data from a relatively large sample of medical school matriculants that can be used as proxy norm tables and cutoff scores for identifying high and low scorers on the JSE-S. The findings have implications for admission decisions under certain conditions, as described above, for identifying those who may need further training to enhance their empathy, in addition to locating the relative standing of a particular individual or a group on the score distribution of the JSE-S.

Acknowledgment

Disclosure Statement

The authors have no conflicts of interest to disclose.

References

1 Hojat M: Empathy in Patient Care: Antecedents, Development, Measurement, and Outcomes. New York, Springer, 2007, p 80.
2 Hojat M, Vergare M, Maxwell K, et al: The devil is in the third year: a longitudinal study of erosion of empathy in medical school. Acad Med 2009;84:1182–1191.
3 Hojat M, Mangione S, Nasca TJ: The Jefferson Scale of Physician Empathy: development and preliminary psychometric data. Educ Psychol Meas 2001;61:349–365.
4 Hojat M, Gonnella JS, Nasca TJ, et al: Physician empathy: definition, components, measurement and relationship to gender and specialty. Am J Psychiatry 2002;159:1563–1569.
5 Hojat M, Gonnella JS, Mangione S, et al: Empathy in medical students as related to academic performance, clinical competence, gender, and academic performance. Med Educ 2002;36:522–527.
6 Hojat M, LaNoue M: Exploration and confirmation of the latent variable structure of the Jefferson Scale of Empathy. Int J Med Educ 2014;5:73–81.
7 Hojat M, Erdmann JB, Gonnella JS: Personality assessment and outcomes in medical education and the practice of medicine: AMEE Guide No. 79. Med Teach 2013;35:e1267–e1301.
8 Hojat M, Paskin DL, Callahan CA: Components of postgraduate competence: analyses of 30 years of longitudinal data. Med Educ 2007;41:282–289.
9 Cohen J: Statistical Power Analysis in Behavioral Sciences. Hillsdale, Erlbaum, 1987.

10 Hojat H, Xu G: A visitor's guide to effect sizes: statistical significance versus practical (clinical) importance of research findings. Adv Health Sci Educ 2004;9:241–249.

11 Blanca MJ, Arnau J, Lopez-Montiel D, et al: Skewness and kurtosis in real data samples. Methodol Eur J Methods Behav Soc Sci 2013; 9:78–84.

12 Hojat M, Mangione S, Nasca TJ: An empirical study of decline in empathy in medical school. Med Educ 2004;38:934–941.

13 Chen D, Lew R, Hershman W, et al: Cross-sectional measurement of medical student empathy. J Gen Intern Med 2007;22:1434–1438.

14 Newton BW, Barber L, Clardy J: Is there hardening of heart during medical school? Acad Med 2008;83:244–249.

15 Ward J, Cody J, Schaal M, et al: The empathy enigma: an empirical study of decline in empathy among undergraduate nursing students. J Prof Nurs 2012;28:34–40.

16 Hojat M, Axelrod D, Spandorfer J, et al: Enhancing and sustaining empathy in medical students. Med Teach 2013;35:996–1001.

17 Van Winkle LJ, Fjortoft N, Hojat M: Impact of a workshop about aging on the empathy scores of pharmacy and medical students. Am J Pharm Educ 2012;76:1–6.

18 Forstater AT, Chauhan N, Allen A, et al: An emergency department shadowing experiences for emergency medicine residents: can it prevent erosion of empathy? Acad Emergency Med 2011;18:s2.

19 Misra-Hebert AD, Isaacson JH, Kohn M, et al: Improving empathy of physicians through guided reflective writing. Int J Med Educ 2012;3:71–77.

20 Magee M, Hojat M: Rocking chair and empathy (letter to the editor). Fam Med 2010;7: 465–466.

21 Hojat M, Louis DZ, Markham FW, et al: Physicians' empathy and clinical outcomes in diabetic patients. Acad Med 2011;86:359–364.

22 Del Canale S, Louis DZ, Maio V, et al: Physicians' empathy and disease complications: an empirical study of primary care physicians and their diabetic patients in Parma, Italy. Acad Med 2012;87:1243–1249.