

The Jefferson Scale of Physician Empathy: Further Psychometric Data and Differences by Gender and Specialty at Item Level

HOJAT, MOHAMMADREZA; GONNELLA, JOSEPH S.; NASCA, THOMAS J.; MANGIONE, SALVATORE; VELOSKI, J. JON; MAGEE, MICHAEL

It is as important to know what kind of a man has the disease, as it is to know what kind of disease has the man. - —Sir William Osler

Researchers agree that empathy has a positive role in clinical outcomes^{2,3,4} and in improving interpersonal relationships,² but they are divided on its definition and components. In the context of health care, we define empathy as “a cognitive (as opposed to affective) attribute that involves an understanding of the inner experiences and perspectives of the patient, combined with a capability to communicate this understanding to the patient.” With the exception of the affective domain, this definition is similar to the conceptualization of empathy by Feighny and colleagues.⁵ The key feature of empathy, according to our definition, is understanding, rather than affective involvement with patients' experiences. The affective domain is a key component of sympathy, rather than empathy.

The Association of American Medical College's Medical School Objectives Project (MSOP)⁶ lists empathy among the educational objectives by emphasizing that medical schools should strive to educate altruistic physicians who are “compassionate and empathetic in caring for patients” and who can understand a patient's perspective by demonstration of empathy.^{6p.13}

Medical educators concede that empathy is a significant factor in patient care that must be cultivated during medical education and can be assessed at admission to medical school.⁷ Likewise, empathy is an important component of “professionalism” in medical practice. Yet, empirical research on empathy among medical students and physicians is scarce. One reason for this dearth of empirical research is the absence of a psychometrically sound and specific research instrument.

A few empathy scales for the general population exist that we previously described^{8,9} but to the best of our knowledge there is no psychometrically sound tool available for measuring empathy among medical students and physicians. There is a need for an operational measure of empathy for medical students and physicians. Such a measure can be used to evaluate the effectiveness of educational interventions aimed at promoting empathy.⁵ In response to this need, we developed the Jefferson Scale of Physician Empathy.^{8,9}

In our previous studies with students, we found that total empathy scores were significantly associated with clinical competence ratings in medical school, but not with licensing examination scores.¹⁰ A significant overlap between empathy and clinical competence constitutes key validity evidence for the empathy scale. In another study, we noticed a significant decline in mean empathy scores during the third year of medical school.¹¹ Such a decline was also observed among internal medicine residents, but it did not reach the conventional level of statistical significance.¹² Overall, we found that female students and physicians scored higher in empathy than males.^{8,9,10}

In our studies with physicians, we noted that physicians in “patient-oriented” specialties obtained a significantly higher average empathy score than those in “technology-oriented” specialties.¹³ Psychiatrists

obtained the highest mean empathy score and anesthesiologists, orthopedists, neurosurgeons, and radiologists received the lowest.⁹ Although we found no significant difference in the total empathy scores between physicians and nurses, the two groups differed significantly on some items.¹⁴ Some of these findings that were consistent with our expectations can be considered as evidence in support of the validity of the empathy scale.

This study was designed to further examine the psychometric properties of the Jefferson Scale of Physician Empathy, and to investigate differences on individual items between men and women and between physicians in specialty areas defined as “people-oriented” and “technology-oriented.”

Method

Participants. Participants included 704 physicians (74% men, 26% women) in the Jefferson Health System affiliated with Thomas Jefferson University Hospital and Jefferson Medical College in the greater Philadelphia region. Participants' mean age was 46.8 years, with a standard deviation of 10.5, range from 29 to 87 years.

Instrument. The Jefferson Scale of Physician Empathy⁹ was used. This scale was originally developed to measure the orientation of medical students toward physician empathy in patient-care situations (Student or S Version).⁸ The scale was constructed based on an extensive review of the literature, followed by pilot studies with samples of physicians, students, and residents.⁸ After several refinements, the instrument included 20 Likert-type items answered on a seven-point scale (1 = “strongly disagree,” 7 = “strongly agree”).

Psychometric data in support of the construct validity and criterion-related validity (convergent and discriminate) of the S-Version of the scale have been reported.⁸ Internal consistency reliability (coefficient alpha) of this version was .89 for medical students and .87 for medical residents.⁸

We also developed a revised version of the scale for physicians and health professionals (Health Professional or HP-Version).⁹ In this version, the wording of the S-Version was modified slightly to make the contents more relevant to the caregiver's empathetic behavior rather than to the student's empathetic orientation or attitudes.

For example, the following item appeared in the S-Version: “Because people are different, it is almost impossible for physicians to see things from their patients' perspectives.” In the HP-Version this item read: “Because people are different, it is almost impossible for me to see things from my patients' perspectives.” These modifications were also intended to make the scale applicable to health care providers other than physicians.

In the S-Version, only three negatively worded items appeared. Negatively worded items are usually used in psychological tests to decrease the confounding effect of the “acquiescence response style” (e.g., the tendency to constantly agree or disagree: yea-, naysayers).¹⁵ In the HP-Version, a balance was maintained by making ten items positively and ten negatively worded. The alpha reliability of the HP-Version was .81. Test—retest reliability was .65 with approximately three to four months between testings.⁹ (Copies of the scales can be obtained from the authors.)

Procedures. The empathy scale, accompanied by a cover letter personally signed by one of the authors (TJN) to increase cooperation, was mailed to 1,007 physicians. The respondents were instructed not to identify themselves, and were assured of strict confidentiality. Two follow-up reminders at four- and eight-week intervals after the original mailing yielded a total of 704 completed surveys, representing a 70% response rate.

Respondents were divided into two groups of “people-oriented” specialties ($n = 462$, 66% of the total respondents who were in primary care specialties [such as family medicine, internal medicine, and pediatrics], obstetrics and gynecology, emergency medicine, psychiatry, and medical subspecialties) and “technology-oriented” specialties ($n = 242$, 34% of the respondents who were in hospital-based specialties [such as anesthesiology, radiology and pathology], surgery and surgical subspecialties). The classification of people- and technology-oriented specialties is common in medical education research.¹⁶

Statistical analyses. Partial item—total score correlations were calculated to identify items with the largest correlations. In addition to analyses of variance (ANOVA) and covariance (ANCOVA), two sets of multivariate analysis of variance (MANOVA) were used (gender was the independent variable in one set and specialty in another), followed by univariate ANOVA to test the significance of differences on each item of the empathy scale. The scores on 20 items of the scale were the dependent variables in both MANOVA models. The effect-size estimates¹⁷ were also calculated for each of the mean differences to detect the practical (clinical) importances of the statistically significant findings.

Results and Discussion

Psychometrics of the items. The mean item scores ranged from a low of 4.8 to a high of 6.5 on the seven-point scale. These findings indicate that responses tend to be skewed toward the upper end of the scale, although physicians actually used the full range of possible responses on all items. The two items with the highest mean score ($M = 6.5$, both were reverse scored) were: “My understanding of how my patients and their families feel is an irrelevant factor in medical treatment,” and “I believe that emotion has no place in the treatment of medical illness.” The item with the lowest mean score ($M = 4.8$) was: “I try to think like my patients in order to render better care.” The standard deviations for the items ranged from 0.9 to 1.6.

The item—total score correlations (Table 1) were all positive and ranged from a low of .30 for two items: “I do not enjoy reading nonmedical literature” (reverse scored) and “My understanding of how my patients and their families feel is an irrelevant factor in medical treatment” (reversed scored), to a high of .6 for two items: “I try to imagine myself in my patients' shoes when providing care to them,” and “My understanding of my patients' feelings gives a sense of validation that is therapeutic in its own right.” All of the item—total score correlations were highly significant ($p < .01$). These findings reaffirm the direction of scoring (indicated by positive correlations) and the significant contribution of each item to the total score of the empathy scale (indicated by significant correlations).

Table 1 shows the item—total score correlations and effect-size estimates of the differences between men and women and between physicians in “people-oriented” and “technology-oriented” specialties.

Gender differences. Statistically significant differences were observed on six of the 20 items of the empathy scale (Wilks' lambda = .94, related multivariate $F(20,644) = 2.21$, $p < .01$). In these items women consistently scored higher than men, confirming our previous findings of gender differences.^{8,9,10} The largest gender-effect size estimates were found for the following items: “My patients feel better when I understand their feelings” ($d = .25$), “I believe that empathy is an important therapeutic factor in medical

treatment" ($d = .21$), and "I consider understanding my patients' body language as important as verbal communication in caregiver—patient relationships" ($d = .21$). According to the operational definitions suggested by Cohen¹⁷ effect-size estimates around .20 are small and negligible, those around .50 are moderate, and those around .80 are large. Therefore, the gender differences are of little practical importance despite their statistical significance.

Five of the six items on which women outscored men were among the items included in the "perspective taking" construct (a core ingredient of empathy), which emerged in a factor analytic study of the empathy scale.⁹ These findings suggest that gender differences are more pronounced on the "perspective taking" aspect of physician empathy.

Specialty differences. Statistically significant differences were observed between "people-oriented" and "technology-oriented" specialties on 11 of the 20 items of the empathy scale (Wilks' $\lambda = .94$, related multivariate $F(20,661) = 2.25$, $p < .01$). Physicians in "people-oriented" specialties consistently outscored their counterparts in "technology-oriented" specialties in all items, providing further evidence in support of our previous findings on the total scale scores.¹³ The largest effect size ($d = .41$) was found for the following item: "An important component of the relationship with my patients in my understanding of the emotional status of themselves and their families." These findings remained unchanged when we controlled for gender using ANCOVA. Five of the 11 items on which physicians in "people-oriented" specialties scored higher than those in "technology-oriented" specialties were among the "perspective taking" component of the empathy scale. Another five items were among the "compassionate care," and one item was among the "standing in the patient's shoes" components of the empathy scale that emerged in a factor-analytic study.⁹

Conclusions and Implementation

Findings of this study showed that all items included in the Jefferson Scale of Physician Empathy are relevant to the operational measure of empathy, although the skewed distribution of scores at the item level suggest that some may need further editing. The significant differences between men and women, and between physicians in "people-oriented" and "technology-oriented" specialties suggest that particular aspects of empathy may be more related to gender and specialty.

It is important to emphasize that the statistically significant differences among physicians do not necessarily indicate a deficiency in empathy in a low-scoring group, for two reasons. First, none of the effect-size estimates is large enough to indicate that the statistically significant difference is clinically important or is out of normal range.

Second, duties involved in the "technology-oriented" specialties do not demand the degree of empathy that is required in the "people-oriented" specialties. Understanding the experiential and emotional status of patients (items with the largest effect-size estimates in Table 1) is more important in primary care than in hospital-based specialties. These findings are valuable in increasing our understanding of similarities and differences among physicians in different aspects of empathy.

In recent years, the patient—physician relationship has become severely strained by changes in the economics of medical practice.¹⁸ Therefore, it is timely and important to study how empathy can be cultivated and assessed among students and physicians and how it contributes to patient outcomes. It is equally important to know what aspects of empathy are more associated with physicians' demographic and career interests.

Our findings suggest that our empathy scale is a reliable and valid instrument for studying physician empathy. We are conducting a multicultural study to examine further the validity of the empathy scale by correlating its scores with patients' perceptions of their physicians' empathy. Once sufficiently validated, the scale could be used in assessing the empathy of individual learners and physicians, thereby meeting a need identified in a recent study on evaluating professional behavior.¹⁹

References

1. Osler W. *Aequanimitas with other Address to Medical Students, Nurses, and Practitioners of Medicine*. 3rd ed. Philadelphia, PA: Blakiston, 1932.
2. Spiro HM, Mccrea Curren MG, Peschel E, St. James D. *Empathy and the Practice of Medicine: Beyond Pills and the Scalpel*. New Haven, CT: Yale University Press, 1993.
3. Nightingale SD, Yarnold PR, Greenberg MS. Sympathy, empathy, and physician resource utilization. *J Gen Intern Med*. 1991;6:420–3.
4. Olsen JK. Relationship between nurse-expressed empathy, patient-perceived empathy and patient distress. *Image: J Nursing Scholarship*. 1996;27:317–22.
5. Feighny KM, Arnold L, Monaco M, Munro S, Earl B. In pursuit of empathy and its relation to physician communication skills: multidimensional empathy training for medical students. *Ann Behav Sci Med Educ*. 1998;5:13–21.
6. AAMC Medical School Objectives Project, (<http://www.aamc.org/meded/msop/>).
7. Kupler DJ, Drew FL, Curtis EK, Rubinstein DN. Personality style and empathy in medical schools. *J Med Educ*. 1978;53:507–9.
8. Hojat M, Mangione S, Nasca TJ, et al. The Jefferson scale of empathy: development and preliminary psychometric data. *Educ Psychol Meas*. 2001;61:349–65.
9. Hojat M, Gonnella JS, Nasca TJ, Mangione S, Vergare M, Magee M. Physician empathy: definition, measurement and relationship to gender and specialty. *Am J Psychiatry*. 2002;159:1563–9.
10. Hojat M, Gonnella JS, Mangione S, et al. Empathy in medical students as related to academic performance, clinical competence, and gender. *Med Educ*. 2002;36:1–6.
11. Hojat M, Mangione S, Nasca TJ, et al. An empirical study of changes in empathy scores among third-year medical students [unpublished manuscript, submitted for publication].
12. Mangione S, Kane GC, Caruso JW, Gonnella JS, Nasca TJ, Hojat M. Assessment of empathy in different years of internal medicine residency. *Med Teach*. In press.
13. Hojat M, Mangione S, Gonnella JS, Nasca TJ, Veloski JJ, Kane G. Empathy in medical education and patient care [letter to the editor]. *Acad Med*. 2001;76:609.

14. Fields SK, Hojat M, Kane G, Mangione S. Comparison of nurses and physicians on empathy. Paper presented at the Isabel Maitland Stewart Research Conference, Teachers College, Columbia University, New York, April 2001.
15. Anastasi A. Psychological Testing. New York; Macmillan, 1976.
16. Lieu TA, Schroeder SA, Altman DF. Specialty choice at one medical school: recent trends and analysis of predictive factors. *Acad Med.* 1989;64:622–9.
17. Cohen J. Statistical Power Analysis for Behavioral Sciences. Hillsdale, NJ: Erlbaum, 1987.
18. Glass R. The patient—physician relationships: JAMA focuses on the center of medicine. *JAMA.* 1996;275:147–8.
19. Arnold L. Assessing professional behavior: yesterday, today, and tomorrow. *Acad Med.* 2002;77:28–41.