

ETHICS AND HUMANITIES

Teaching Empathy to First Year Medical Students: Evaluation of an Elective Literature and Medicine Course

JOHANNA SHAPIRO, ELIZABETH H. MORRISON & JOHN R. BOKER

Department of Family Medicine, University of California Irvine College of Medicine, USA

ABSTRACT **Background:** *Empathy is critical to the development of professionalism in medical students, and the humanities—particularly literature—have been touted as an effective tool for increasing student empathy. This quantitative/qualitative study was undertaken to assess whether reading and discussing poetry and prose related to patients and doctors could significantly increase medical student empathy and appreciation of the relevance of the humanities for their own professional development.*

Method: *In 2000–2001, first year students (n=22) volunteered for an eight-session literature and medicine elective and were randomly assigned to either immediate participation in the class or a wait-list group, who participated in the same class 6 months later. Complete pre- and post-intervention data for 16 students from both groups were obtained for two quantitative measures of empathy and an attitudes-toward-the-humanities scale. Students also participated in a qualitative group interview pre- and post-intervention.*

Results: *Empathy and attitudes toward the humanities improved significantly ($p < 0.01$) after participation in the class when both groups of students were combined. The scaled treatment effect size was in the moderate range (≥ 0.60 standard deviation units) for both measures that had statistically significant pre-to-post changes. Furthermore, student understanding of the patient's perspective became more detailed and complex after the intervention. Students were also more likely post-intervention to note ways reading literature could help them cope with training-related stress.*

Conclusion: *A brief literature-based course can contribute to greater student empathy and appreciation for the value of humanities in medical education.*

KEYWORDS *Medical humanities, medical education, empathy.*

Author for correspondence: Johanna Shapiro, Ph.D., Dept. of Family Medicine, 101 City Drive South, Rte 81, Bldg 200, Orange, CA 92868, USA. Tel: 949-824-3748. Fax: 714-456-7984. E-mail: jfshapir@uci.edu

Introduction

Professionalism in medical education must include the development of empathy (Gianakos, 1996; Marcus, 1999), the capacity to participate deeply in another's experience (Spiro, 1992). However, although learner empathy has been linked not only to patient satisfaction (Smith *et al.*, 1995), but to clinical competence (Hojat *et al.*, 2002), evidence suggests that empathy actually declines over the course of undergraduate medical education (Newton *et al.*, 2000; Lu, 1995). To date, most pedagogical efforts have approached empathy as a set of cognitive and behavioural skills (Platt & Keller, 1994; Burack *et al.*, 1999; Winefield & Chur-Hansen, 2000). Concerns have been expressed about whether this instructional method is sufficient to produce truly empathic physicians (Henry-Tillman *et al.*, 2002).

If empathy is as much art as science (Misch, 2002), then perhaps it may be developed through studying the humanities (Hunter *et al.*, 1995; Charon, 2000). From a theoretical perspective, the humanities engage the emotions as well as the intellect, thereby achieving deep understanding of and insight into the experience of another (Charon, 2001). Such comprehensive empathy is hypothesized to translate more powerfully into clinical situations than that achieved by cognitive-behavioural approaches. However, although learner response to medical humanities is generally positive (Horowitz, 1996; Anderson, 1998), we still know little about what the humanities can contribute specifically to the encouragement of empathy in medical students and to their professionalism.

At our home institution, training in empathy is provided to pre-clinical medical students in a cognitive-behavioural format as part of a communication skills course. We wanted to explore supplemental pedagogic methods that might positively influence student empathy. Specifically, we tested the hypothesis that reading and discussing literature about patients' experiences of illness and the doctor-patient relationship could significantly improve medical student empathy and appreciation of the relevance of the humanities for professional development.

Method

Study Design

This study employed a qualitative/quantitative, repeated-measures design where the primary aim was to measure the effects of a literature and medicine elective on student empathy and attitudes. Students who volunteered for the study were randomly assigned to either an immediate intervention group (Group 1: participation in the literature and medicine course) or a wait-list group (Group 2: delayed intervention) by a research assistant using a random number table (Figure 1).

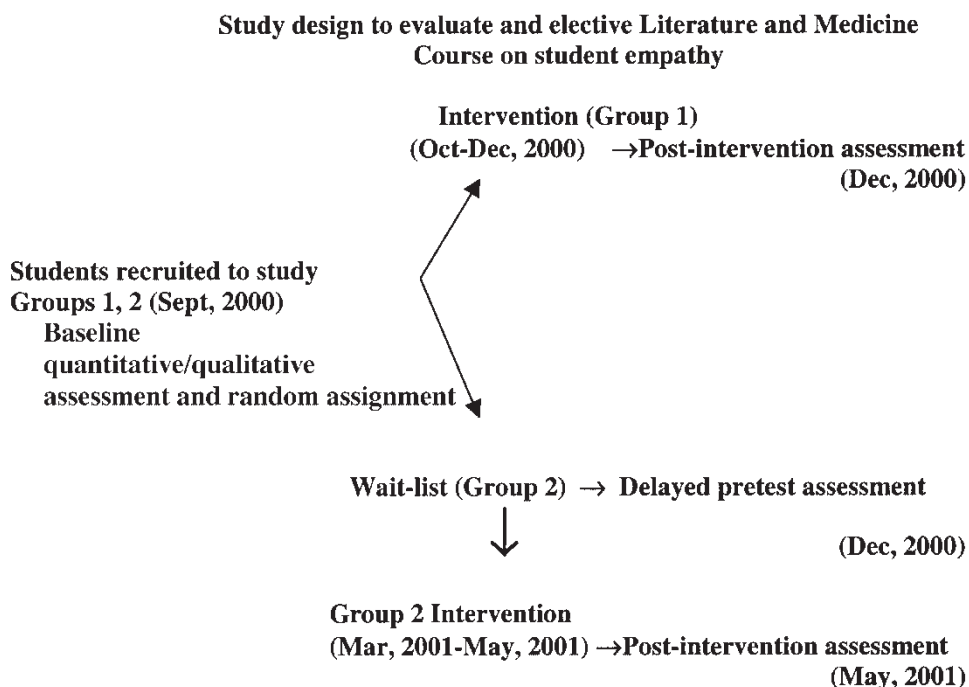


Figure 1. Study design to evaluate an elective literature and medicine course on student empathy.

This modified cohort-control design (McKillip, 1979) allowed the collection of pre- and post-intervention measurements on two separate groups of students, where the intervention for Group 2 lagged behind that for Group 1. This design controlled for the effects of exposure to the standard curriculum and the passage of time on the dependent variables. Since wait-list students eventually received the intervention (after 6 months of exposure to standard curriculum), we replicated the data collection and intervention in two separate groups.

Subjects and Procedures

All first year students ($n = 92$) at a public, allopathic medical school in Southern California were invited by e-mail, flier and personal announcement to enroll in a literature and medicine elective and participate in a study of its effects on student attitudes. Twenty-two first year students volunteered (24% of the class).

Intervention

The educational intervention consisted of eight small-group reading and discussion sessions, offered as a longitudinal curriculum for 1 hour twice monthly for a total of 8 hours of teaching. The class was co-taught by a PhD psychologist and a rotating primary care physician faculty member.

The format of the class consisted of on-site readings of poetry, skits and short stories that addressed the following topics: doctor-patient relationship, physical examination, listening to patients, pain, sexuality, cross-cultural issues, lifestyle modification/noncompliance and geriatrics. We placed special emphasis on understanding of, and identifying with, different points of view in the texts, including those of physicians, patients, and family members, as well as their own.

Assessment

All 22 participating students provided baseline data (September 2000) that included a group interview, administration of two quantitative measures of empathy, and one instrument assessing attitudes toward the humanities as a tool for professional development. The group interview was conducted by the first author following focus group guidelines (Krueger, 1998a). All group interviews were tape-recorded, and key portions transcribed. For Group 1, this comprised their pre-intervention data.

At the conclusion of the Group 1 intervention (December 2000), we collected their post-intervention data by conducting a follow-up interview and repeating the administration of the empathy/attitude measures. Simultaneously, we also repeated the quantitative measures for Group 2 as a delayed pretest. After Group 2 completed the intervention (May 2001), we gathered their post-intervention data with a follow-up interview and re-administration of the quantitative measures a final time to this group only.

Measures

The first empathy measure was a 20-item version of the Empathy Construct Rating Scale (ECRS) (LaMonica, 1981, 1996) scored on a 6-point rating scale (1 = extremely unlike self, ... 6 = extremely like self). The ECRS has been demonstrated to have high reliability and discriminant validity. Typical items include self-assessments of ability to listen carefully, accurately paraphrase the feelings of others, and checks to see if one's understanding of another's experience is valid.

We also employed the Balanced Emotional Empathy Scale (BEES) (Mehrabian *et al.*, 1988; Mehrabian, 1996), a reliable and valid 30-item instrument scored on a 9-point rating scale (+4 = very strong agreement, 0 = neither agreement nor disagreement, and -4 = very strong disagreement). The BEES probes the extent to which the respondent can feel others' suffering or take pleasure in their happiness.

Finally, we developed a 9-item attitude-toward-the-humanities measure that used a 6-point rating scale¹ (where higher rating values were indicative of more positive attitudes) to assess the extent to which students thought the humanities could be useful in their professional development (Rucker & Shapiro, 2003). Representative items include: "I am likely to turn to the humanities to help me understand my experience in medical school." "The humanities are a useful

tool to explore personal feelings evoked by illness experiences and the doctor-patient relationship.” To assess the internal consistency reliability of each measure, we calculated Cronbach’s coefficient alpha, using the data from all students who completed each separate measure.

The question route was based on procedures for generating focus group discussion questions (Krueger, 1998b). We asked these questions: (1) What is empathy? (2) How can study of the humanities improve understanding of the patient’s perspective? (3) How can the humanities make you a better physician? (4) How can the humanities help students cope with the experience of medical school? This question route was pilot-tested for clarity and comprehensibility with three second-year medical students not otherwise connected with the elective.

Data Analysis

Because of the relatively small sample size of each group in the cohort-control design, we looked for ways to increase the power of our statistical analysis. After first determining that no baseline differences existed between the groups on any of the three scales (Mann–Whitney test) and then that no differences existed in Group 2’s baseline and delayed pretest scale scores (Wilcoxon signed rank test), we combined students into a single group. Group 2’s baseline and delayed pre-test scores were averaged, resulting in a single score for each scale that was subsequently treated as the pre-intervention data for Group 2 students.

The main analysis of the quantitative data used the Wilcoxon signed rank test and was a comparison of pre- and post-intervention scores on each scale only for the 16 students who had complete sets of pre- and post-data. Students having missing data for any measure at any administration were deleted from the main analyses. We made inferences using nonparametric statistics because of the study’s small sample size and to avoid making untenable assumptions about the underlying distribution of scores. The two-tailed alpha level was set at $p < 0.05$, and the nonparametric tests were performed with standard statistical software (SPSS, 1996). We did not adjust the criterion for statistical significance for multiple comparisons.

The scaled treatment effect sizes were calculated by the formula: $(M2-M1)/SD1$, where $M1$ and $M2$ are the respective group means at pre- and post-intervention, and $SD1$ is the pooled pre-intervention standard deviation. Behavioural science researchers generally define an effect size of 0.20 as small, 0.50 as moderate, and 0.80 as large. Effect size provides insight about the practical or educational significance of change measured before and after an intervention.

The data from the three interview groups were interpreted using a content analysis (Marshall & Rossman, 1989; Strauss & Corbin, 1991; Crabtree & Miller, 1999). The first author and a research assistant listened to audio-tapes from all groups and made verbatim transcriptions of key portions. The data

were first reviewed to identify key words and phrases, then organized into more inclusive categories, and finally into major themes. The second author reviewed and modified these conclusions. Member checking (Lincoln & Guba, 1985) was conducted by the first author between 1 and 3 months post-data collection. Six students were contacted, three from Group 1 and three from Group 2. The major qualitative conclusions were presented, and students generally agreed they represented the substance of the group discussions.

Results

Sixteen of the 22 volunteer students (73%) completed all pre- and post-quantitative measures. All 22 students participated in the first focus group. In Group 1, 10 students (91%) participated in their follow-up focus group and also completed all pre- and post-quantitative measures. For Group 2, nine students (82%) participated in their follow-up interview, and six of them (55%) completed all quantitative measures.

Students were approximately evenly distributed between male and female (female $n=12$; 55%). Their age range was 21–30 years (mean = 23.4 [1.9]). Students self-identified themselves as non-Hispanic white ($n=6$; 27%); Asian ($n=8$; 36%); and other ($n=8$; 36%). Of the students, 19 (86%) had general science majors. Eight (36%) thought they would enter primary care specialties, another eight (36%) anticipated applying to other specialties and six (27%) were unsure of their specialty choice. This profile does not differ substantially from that of students enrolled in the entire class (i.e., 52% female, mean age = 23.5, 34% non-Hispanic white, and 35% indicating an intention to choose primary care specialties).

The rating scales yielded highly reliable scores on all administrations. The mean coefficient-alpha reliability for the ECRS in this study was 0.84; for the BEES was 0.81; and for the attitude scale was 0.78.

Both Group 1 and Group 2 were identical at baseline on the three quantitative measures, indicating the validity of the randomization process and the equivalency of the two groups (Mann–Whitney $p=0.62$, 0.09, and 0.66 for ECRS, the attitudes-toward-the-humanities, and the BEES scale scores, respectively). Also, analysis of Group 2's scores yielded no significant within-group differences between baseline and delayed pretest scores (Wilcoxon $p=0.84$, 0.33, and 0.10, for ECRS, attitudes-toward-the-humanities, and BEES scale scores, respectively). The latter finding suggests the absence of major history and maturational effects on the study outcomes and also justifies the aggregation of the scale scores from baseline and delayed pretest administrations into a single set of pre-intervention scores for Group 2.

There was no significant pre-to-post difference on average ECRS scores when Groups 1 and 2 were combined into a single analysis (pre = 92.3, post = 94.6; $p=0.27$). However, statistically significant pre-to-post increased

average scores on both the BEES (pre = 57.0, post = 68.9; $p < 0.01$) and the attitudes-toward-the-humanities scale (pre = 43.5, post = 46.0; $p < 0.01$) were identified. Scaled treatment effect sizes were in the moderate range (≥ 0.60 standard deviation units) for both the BEES (20% improvement over pre-test) and the attitudes-toward-the-humanities scale (6% improvement) (Table 1).

The Pearson product-moment correlation for the two empathy measures at pre-intervention was low and not significant ($r = 0.14$), suggesting that they measured different dimensions of empathy. The correlation between empathy measures was higher post-intervention, but did not reach statistical significance ($r = 0.48$, $p = 0.06$). The attitude scale did not correlate significantly with empathy either at pre-intervention ($r = 0.16$ with ECRS, and $r = -0.19$ with BEES) or post-intervention ($r = 0.30$ and 0.31 respectively). Female students, Asian students and students planning to enter primary care showed significantly more empathy post-intervention as measured by the BEES scale. In these three demographic groups, average pre-post changes in BEES scores were associated with scaled treatment effect sizes of 0.68 for females, 0.91 for Asians and 1.39 for primary care respectively, all bordering at or falling above the widely accepted definition of a large effect. Female students improved by 24% over baseline, Asian students by 18% and those choosing primary care by 54%. In terms of attitudes-toward-the-humanities, there were no significant pre-post changes (Table 2).

Qualitative analysis of the combined pre- and post-intervention group interview data suggested that students' definitions of empathy remained essentially unchanged. However, student understanding of the patient's perspective became more detailed and complex post-intervention. For

Table 1. Results from comparing pre-post differences on three quantitative measures for 16 medical students receiving a teaching intervention on empathy, University of California, Irvine, College of Medicine, 2000–2001

Measure	Pre-test Mean (s.d.)	Post-test Mean (s.d.)	Post – Pre	
			Z*	p
ECRS [†]	92.3 (8.2)	94.6 (8.9)	– 1.10	0.27
Attitude scale [‡]	43.5 (4.1)	46.0 (4.3)	– 2.54	< 0.01
BEES [§]	57.0 (20.2)	68.9 (17.5)	– 2.53	< 0.01

* Wilcoxon signed rank test was used for within-group, pre-post comparisons

[†]ECRS is a 20-item version of the Empathy Construct Rating Scale scored on a 6-point rating scale, where 1 = extremely unlike self and 6 = extremely like self.

[‡]Attitude scale is a 9-item attitude-toward-the-humanities measure scored on a 6-point rating scale, where 1 = low endorsement and 6 = high endorsement.

[§]BEES is a 30-item Balanced Emotional Empathy Scale scored on a 9-point rating scale, where +4 = very strong agreement, 0 = neither agreement nor disagreement, and –4 = very strong disagreement.

Table 2. Results from comparing pre-post differences by gender, ethnicity, and future medical specialty on the Balanced Emotional Empathy Scale (BEES)* for medical students receiving a teaching intervention on empathy, University of California, Irvine, College of Medicine, 2000–2001

	<i>n</i>	Pre-test Mean (SD)	Post-test Mean (SD)	Post - Pre	
				<i>Z</i> [†]	<i>p</i>
Gender					
Male	6	53.8 (19.3)	62.3 (17.9)	– 1.36	0.17
Female	10	59.0 (21.4)	72.9 (16.9)	– 2.19	0.03
Ethnicity					
White	5	58.7 (24.6)	72.4 (15.9)	– 0.94	0.44
Asian	7	55.2 (17.7)	65.1 (20.8)	– 2.03	0.04
Other	4	58.1 (24.2)	71.3 (16.4)	– 1.83	0.13
Specialty					
Primary care	6	50.3 (20.7)	77.3 (11.6)	– 2.21	0.03
Other specialty	6	65.4 (19.9)	66.0 (19.9)	– 0.31	0.84
Unknown	4	54.6 (20.5)	60.8 (20.0)	– 1.46	0.25

*BEES is a 30-item measure scored on a 9-point rating scale, where +4 = very strong agreement, 0 = neither agreement nor disagreement, and –4 = very strong disagreement.

[†]Wilcoxon signed rank test was used for all within-group, pre-post comparisons.

example, typical pre-intervention comments included “Reading literature will help me get insight into the human condition.” “Studying poetry will help me understand other people’s feelings.” Typical post-intervention comments are illustrated by the following: “The poems and stuff helped me see how disease affects patients’ daily lives.” “After this course, I realize patients can be afraid, belligerent, and vulnerable”.

Post-intervention insights about how literature could make students better physicians were also more specific and sophisticated. Again, typical pre-intervention comments were “Literature might help me be more insightful about patients” and “The humanities will make me become more well-rounded.” The following were more representative of post-intervention comments: “I’ve learned about how to change my perspective from myself to looking at a situation from the patient’s point of view.” “By reading, I’ve learned to listen for what the patient needs and hopes for.”

Finally, pre-intervention students did not express many thoughts about how reading literature could help them cope with medical school. Post-intervention, students were more positive about literature’s potential usefulness in dealing with training-related stresses. One student commented, “I’ve learned the importance of expressing your feelings with your classmates.” Another student spoke of “reading as a refuge, or a source of solutions,” while another said, “This class is like going to church—it reminds me of the big picture.”

Discussion

This study had several limitations, particularly constraints on generalizability, including the small number of self-selected subjects and the fact that subjects were recruited from a single institution. Because we were unable to administer measures to the whole class, it is possible that students enrolled in the elective were more empathic than their non-participating classmates. In addition, we did not attempt a long-term follow-up of these students as they progressed through training. That would have enabled us to link self-perceptions of empathy to actual student behaviour in clinical situations, as well as to patient evaluations, surely the ultimate outcomes of interest. Finally, a problem with missing data reduced the statistical power and forced us to use less powerful nonparametric inferential tests.

Despite these limitations, the study found significant improvements in medical student empathy and attitudes toward the humanities after participation in a literature-based intervention. Statistically significant pre-post changes were accompanied by scaled treatment effect sizes that were moderate by the conventional definition. Pre-post effect sizes were even greater for some of the demographic variables. Qualitative data confirmed that, after intervention, students had a more detailed and sophisticated understanding of how studying literature could help them understand their patients better and to become better physicians. Finally, both quantitative and qualitative data indicated that, at the conclusion of the intervention, students were more likely to turn to literature to help them cope with the stresses of professional life.

Based on the fact that empathy scores in the delayed intervention group did not increase as a function of history and maturation alone, it is possible to infer that exposure to the standard cognitive-behavioural method of instruction did not appear to improve empathy, at least according to student self-report. This finding reflects extant concerns in the professional literature about the most appropriate way to teach empathy (Winefield & Chur-Hansen, 2000; Shapiro, 2002).

A related issue is the differences in student performance on the two empathy measures. Face validity inspection of the ECRS items suggests they measure cognitive/behavioural aspects of empathy, while the BEES items assess the construct's emotional aspects (Arnette, 2003). The intervention appeared more successful in promoting emotional aspects of empathy, suggesting that studying literature may exert more influence on particular dimensions of empathy than on others.

Finally, this study raised the possibility that specific subgroups of medical students may be especially receptive to a literature-based intervention designed to improve empathy. Specifically, although empathy scores improved among all demographic subgroups measured, women students, Asian students and students interested in primary care seemed to respond more strongly to the intervention. Further research is needed to determine whether specific attitudes

or attributes characteristic of these groups predispose them to improving their empathic tendencies, and to identify more effective ways of increasing empathy among all students.

Conclusion

The findings demonstrate that a literature-based intervention can exert a significant and meaningful influence on certain aspects of students' self-reported empathy and positively affect their attitudes toward the humanities as a useful tool in professional development. After a relatively brief exposure to literature, students also had a more sophisticated understanding of patients and were more likely to think of literature as an effective coping mechanism to deal with training-related stress. Further research must determine whether this effect can translate into actual changes in student behaviour that measurably enhances patient well-being.

Acknowledgments

The authors would like to thank the 22 first year students at the University of California Irvine College of Medicine who made this study possible. Also, the contributions of Meha Harthill, PhD, Corollos Abdelshehid and Sharon Brown should be noted for the assistance they provided as research assistants in this project.

This project was supported by funds from HRSA Pre-doctoral Training in Primary Care Grant #HP 000224-02 (Elizabeth Morrison, MD, MEd, principal investigator) and the UCI Department of Family Medicine.

Note

1. Readers may contact the first author for a copy of this scale.

References

- ANDERSON, C.M. (1998). "Forty acres of cotton waiting to be picked": Medical students, story-telling, and the rhetoric of healing. *Literature and Medicine*, 17, 280–297.
- ARNETTE, J.K. (2003). Physicians' empathy: Origin of the deficit. Paper presented at the 24th Annual Meeting & Scientific Session of the Society of Behavioral Medicine.

- BURACK, J.H., IRBY, D.M., CARLINE, J.D., ROOT, R.K. & LARSON, E.B. (1999). Teaching compassion and respect: Attending physicians' responses to problematic behaviors. *Journal of General Internal Medicine*, 14, 49–55.
- CHARON, R. (2000). Literature and medicine: origins and destinies. *Academic Medicine*, 75, 23–27.
- CHARON, R. (2001). Narrative medicine: A model for empathy, reflection, profession, and trust. *JAMA*, 286, 1897–1902.
- CRABTREE, B.F. & MILLER, W.L. (1999). *Doing Qualitative Research*. 2nd edn. Thousand Oaks: Sage Publications, pp. 163–178.
- GIANAKOS, D. (1996) Empathy revisited. *Archives of Internal Medicine*, 156, 135–156.
- HENRY-TILLMAN, R., DELONEY, L.A., SAVIDGE, M., GRAHAM, C.J., & KLIMBERG, V.S. (2002). The medical student as patient navigator as an approach to teaching empathy. *American Journal of Surgery*, 183, 659–662.
- HOJAT, M., GONNELLA, J.S., MANGIONE, S., NASCA, T.J., VELOSKI, J.J., ERDMANN, J.B., CALLAHAN, C.A. & MAGEE, M. (2002). Empathy in medical students as related to academic performance, clinical competence and gender. *Medical Education*, 36, 522–527.
- HOROWITZ, H.W. (1996). Poetry on rounds: A model for the integration of humanities into residency training. *Lancet*, 347, 447–449.
- HUNTER, K.M., CHARON, R. & COULEHAN, J.L. (1995). The study of literature in medical education. *Academic Medicine*, 70, 787–794.
- KRUEGER, R.A. (1998a). *Moderating Focus Groups*. Thousand Oaks: Sage Publications. Pp. 3–36, 45–56.
- KRUEGER, R.A. (1998b). *Developing Questions For Focus Groups*. Thousand Oaks, CA: Sage, Part I, pp. 3–18; Part II, pp. 19–60.
- LAMONICA, E.L. (1981). Construct validity of an empathy instrument. *Research in Nursing and Health*, 4, 389–400.
- LAMONICA, E.L. (1996). *Empathy Construct Rating Scale*. Santa Clara, CA: Xicom, Inc., pp. 1–8.
- LINCOLN, Y.S. & GUBA, E.G. (1985). *Naturalistic Inquiry*. Newbury Park: Sage Publications.
- LU, M.C. (1995). Why it is hard for me to learn compassion as a third year medical student. *Cambridge Quarterly of Health Care Ethics*, 4, 454–458.
- MARCUS, E.R. (1999). Empathy, humanism, and the professionalization process of medical education. *Academic Medicine*, 74, 1211–1215.
- MARSHALL, C. & ROSSMAN, G.B. (1989). *Designing Qualitative Research*. Newbury Park: Sage Publications, pp. 112–120.
- MCKILLIP, J. (1979). Impact evaluation of service programs: Three flexible designs. *Evaluation Quarterly*, 3, 97–104.
- MEHRABIAN, A. (1996). *Manual for the Balanced Emotional Empathy Scale (BEES)*. (Available from Albert Mehrabian, 1130 Alta Mesa Road, Monterey, CA 93940, USA).
- MEHRABIAN, A., YOUNG, A.L. & SATO, S. (1988). Emotional empathy and associated individual differences. *Current Psychology: Research & Review*, 7, 221–224.
- MISCH, D.A. (2002). Evaluating physicians' professionalism and humanism: The case for humanism “connoisseurs”. *Academic Medicine*, 77, 489–495.

- NEWTON, B.W., SAVIDGE, M.A., BARBER, L., CLEVELAND, E., CLARDY, J., BEEMAN, G. & HART, T. (2000). Differences in medical students' empathy. *Academic Medicine*, 75, 1215.
- PLATT, F.W. & KELLER, V.F. (1994). Empathic communication: a teachable and learnable skill. *Journal of General Internal Medicine*, 9, 222–226.
- RUCKER, L. & SHAPIRO, J. (2003). Becoming a doctor: student creative projects in an internal medicine clerkship. *Academic Medicine*, 78, 391–397.
- SHAPIRO, J. (2002). How do physicians teach empathy in the primary care setting? *Academic Medicine*, 77, 323–328.
- SMITH, R.C., LYLES, J.S., METTLER, J.A., MARSHALL, A.A., VAN EGEREN, L.F., STOFFELMAYR, B.E., OSBORN, G.G. & SHEBROE, V. (1995). A strategy for improving patient satisfaction by the intensive training of residents in psychosocial medicine: A controlled, randomized study. *Academic Medicine*, 70, 7297–7232.
- SPIRO, H. (1992). What is empathy and can it be taught? *Annals of Internal Medicine*, 116, 843–846.
- SPSS EXACT TESTS (1996). Chicago, IL: SPSS Inc.
- STRAUSS, A. & CORBIN, J. (1991). *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*. Newbury Park, CA: Sage.
- WINEFIELD, H.R. & CHUR-HANSEN, A. (2000). Evaluating the outcome of communication skill teaching for entry-level medical students: does knowledge of empathy increase? *Medical Education*, 34, 90–94.