

**DEVELOPMENT,  
RELIABILITY, AND  
VALIDITY OF THE  
MEDICAL HELPING  
RELATIONSHIP  
INVENTORY**

**FREDRIC M. WOLF  
ERNESTO REZA  
JAMES O. WOOLLISCROFT  
JUDITH G. CALHOUN  
GRACE JORDISON BOXER  
ANDREW J. ZWEIFLER**  
*University of Michigan Medical School*

*Results of five studies, which were conducted over a 4-year period, provide initial evidence that a new instrument, the Medical Helping Relationship Inventory (MHRI), can reliably and validly be used in the evaluation of medical communication skills. The final version of the MHRI consists of five scales that assess preferences for understanding, probing, focusing, reassuring, and judgmental responses to patients. Coefficient alpha reliability estimates for these scales exceeded estimates obtained for both the counseling Helping Relationship Inventory and a preliminary version of the MHRI and ranged between .62 and .91 (Studies 1 and 2). Evidence of convergent, discriminant, and construct validity is provided in Studies 3 and 4. Results were replicated and cross-validated in Study 5. The authors believe that the MHRI may be useful in other settings (e.g., other medical or professional schools) and with other groups of health professionals (e.g., house officers, community physicians, nurses, social workers, and others).*

**AUTHORS' NOTE:** This study was supported in part by grant # NIH P50 AG 08671 from the National Institute on Aging and by a grant from the University of Michigan Faculty Development Fund, Center for Research on Learning and Teaching. One of the coauthors, Ernesto Reza, is currently on the faculty at California State University at San Bernardino.

**T**he importance of communication skills in fostering effective doctor-patient interpersonal relationships is well recognized. Evidence exists that effective physician communication is related to both patient adherence to treatment regimens and satisfaction with medical care (Haug & Lavin, 1979; Ingelfinger, 1980; Lundberg, 1985; Mechanic, 1985). Physicians' reactions to patients have been called the key to teaching and providing humanistic medicine (Gorlin & Zucker, 1983). Leaders in medical practice, licensure, and education have been reexamining the nature and importance of the interpersonal aspects of care, and certification bodies have been developing guidelines and criteria for competency that include the humanistic practice of medicine (e.g., American Board of Internal Medicine, 1979, 1983; Muller, 1984).

While medical communication and interviewing training programs have become widespread in American medical schools (Kahn, Cohen, & Jason, 1979), precisely defining and evaluating in behavioral terms just what constitutes "good" communication and humanistic qualities has been difficult. In order to develop and implement educational programs, medical educators have borrowed many of the concepts, ideas, and techniques used in counseling and clinical psychology training programs. Examples of this include the use of empathy, open questions, and rapport building in facilitating more positive doctor-patient interactions (DiMatteo, 1979; Nardone, Reuler, & Girard, 1980). Medical education also has borrowed from psychology many of the evaluation instruments developed for assessing the effectiveness of this training. For example, several programs have used the Helping Relationship Inventory (HRI; Jones, 1973) to evaluate medical students' preferences for different types of communication responses to patients (Harris, Eckert, Petzel, & Westermeyer, 1984; McPherson, Sachs, Knopp, & Wolf, 1984; Wolf, Woolliscroft, Calhoun, & Boxer, 1987). These studies have reported significant increases in students' preferences for understanding (or empathic) responses, as well as decreases in preferences for evaluative (or judgmental) responses as a result of training. Other programs have used Carkhuff's (1969) measures of empathic responding to assess the impact of training on medical students (Engler, Saltzman, Walker, & Wolf, 1981; Jarski, Gjerde, Bratton, Brown, & Matthes, 1985; Wolf,

Savickas, Saltzman, & Walker, 1984). These measures, as well as many of the others used in medical training, suffer from the limitation of comprising patient cases that contain content primarily of a psychological or counseling, and not medical, nature. For example, of the four measures of empathy used in a recent study published in the *Journal of Medical Education* (Jarski et al. 1985), none were specifically developed for use in medical education.

Because of the paucity of medically specific instruments, we sought to develop and validate a measure that could be used in both the teaching and evaluation of *medical* communication skills, which was the rationale for our development of the Medical Helping Relationship Inventory (MHRI). The studies reported here were conducted over a 4-year period in an effort to provide a reliable instrument for teaching, evaluation, and research that could be used to assess the preferences of medical students, house officers, physicians, and other health professionals for various types of communication with patients.

## CONCEPTUAL MODEL AND ITEM DEVELOPMENT

### DEFINITIONS AND CONCEPTUAL OVERVIEW

Jones's (1973) Helping Relationship Inventory (HRI) was used as a model for developing the Medical Helping Relationship Inventory (MHRI). Even though this instrument was developed specifically for use with counseling psychology students and other nonmedical individuals, several studies at diverse medical schools have shown it to be sensitive to change in medical students' communication skills as a result of educational intervention (Harris et al., 1984; McPherson et al., 1984; Wolf et al., 1987). The only wording changes made in the version of the HRI used in these medical settings were to replace the words "counselor" and "client" with "physician" and "patient," respectively. No other content changes were made to make this instrument reflect medical, and not counseling, content. The HRI is composed of a series of patient cases, each of which contain a statement that the patient makes to the physician. Respondents to the HRI are asked to assume that they are the patient's physician and are asked to

rank-order their preferences for each of five different statements that could be used to respond to the patient. Each of these responses is representative of a different type of communication skill. An *understanding* response is defined as accurately reflecting the content and feeling of what patients say, how they "see" it. This represents the Rogerian reflection of feeling, or empathic responding approach (Jones, 1973). An *interpretative* response attempts to teach or impart meaning to patients, implying what they might or ought to think. A *probing* response attempt to elicit more information by provoking further discussion along a certain line. A *supportive* response attempts to reassure or pacify patients and often implies that they need not feel that way. An *evaluative* response makes a judgment of the relative goodness, appropriateness, effectiveness, or rightness of what patients say.

#### SCALE CONSTRUCTION AND ITEM GENERATION

To develop medical cases for the MHRI, statements that patients might make to physicians were selected from the literature and actual patient-health provider interactions. A description of the patient content for the 10 cases comprising the final version of the MHRI is provided in Table 1. This content focused on such medically relevant topics as the patient's concern about the appropriateness and effectiveness of therapy, the invasiveness of a diagnostic test, involvement in decision making, and so on. To develop realistic and relevant responses to these 10 patient concerns to be ranked on the MHRI, over 100 medical students and health professionals were asked to respond briefly in one or two sentences to these patient cases as if they were the patient's physician. Another group of health professionals (two physicians, four educational psychologists, and one nurse) experienced in teaching interviewing skills to medical students served as expert judges to classify students' actual open-ended responses to these patient cases according to one of the five types of responses defined earlier for the counseling HRI. Only when all seven judges unanimously agreed on the statement's classification was it considered for use as a response option for a closed-ended MHRI patient case. Using this process, four responses representing understanding, probing, supportive, and evaluative response types, respectively, were

**TABLE 1**  
**Description of the Content of the 10 Patient Cases**  
**Comprising the Medical Helping Relationship Inventory**

- 
- 
1. Concern about appropriateness of therapy
  2. Concern about impending surgery
  3. Concern about pain
  4. Concern about deciding whether to undergo therapy
  5. Concern about invasiveness of diagnostic test
  6. Concern about not being involved in decision making
  7. Concern about impact of illness
  8. Concern about effectiveness of therapy
  9. Concern about length of waiting time to see doctor
  10. Concern about what other people will think about their self-care behavior
- 

selected for each of the patient cases comprising the pilot version of the MHRI. Content validity of the MHRI was enhanced by using this content analysis procedure to select items and develop the scales.

The pilot version of the MHRI asked students to rate the four responses provided in each case on a scale of 1 (*most preferred*) to 4 (*least preferred*). Total scores for each type of communication response were obtained by summing the scores across cases for the items for that category. Lower scores are indicative of greater preference for that type of response to patients. Interpretative response types were not included because of their low reliability with medical students (see Study 1) and the higher level of skill needed to use them appropriately. The final version of the MHRI, however, contains five response options, which are described under the description of the initial validation studies (Studies 3 and 4).

## PRELIMINARY WORK: STUDIES 1 AND 2

### SUBJECTS AND PROCEDURE

Studies 1 and 2 were conducted to examine the reliability of the counseling HRI with medical students (Study 1) and the performance of the preliminary version of the MHRI (Study 2). These studies were undertaken to provide preliminary data for guidance in the develop-

ment of the final version of the MHRI. Study 1 used the standard version of the 10-item HRI, as described in previous studies cited in the literature. Study 2 pilot-tested a three-item version of a new scale, the MHRI, in order to obtain preliminary reliability data to ascertain whether shifting to more medically oriented cases showed promise. The MHRI was patterned after the HRI but contained new patient cases that focused on medical, rather than counseling, content and situations.

Subjects in Study 1 were 156 first-year medical students in the class of 1986; subjects in Study 2 were 154 first-year medical students in the class of 1987. Approximately half of the subjects in each study had received specific training in medical interviewing and communication skills before completing the HRI (Study 1) or MHRI (Study 2).

## RESULTS AND DISCUSSION

Coefficient alpha (Cronbach, 1951) reliability estimates for the various HRI and MHRI scales in Studies 1 and 2 are summarized in Table 2. For the HRI in Study 1, reliability estimates ranged from a low of .39 for the Interpretive scale to a high of .84 for the Understanding scale. Thus the understanding responses exhibited very good internal consistency, the interpretative responses exhibited very low internal consistency, and the evaluative, supportive, and probing responses exhibited moderate levels of internal consistency. Based primarily on the very poor reliability of the Interpretive scale, it was not included in the preliminary MHRI developed and pilot-tested in Study 2.

Results for the pilot version of the MHRI in Study 2 indicated similar, although generally slightly lower, reliability for the Understanding (.82), Probing (.50), Supportive (.58), and Evaluative (.44) scales. These findings of approximately equal reliability for the HRI and the pilot MHRI were encouraging, in as much as the pilot version of the MHRI contained only 3 patient cases, whereas the HRI had 10. Because reliability is known to increase with test length (Stanley & Hopkins, 1972), we believed that the final MHRI would perform even more reliably once it was revised and lengthened.

**TABLE 2**  
**Coefficient Alpha Reliability Estimates for the Counseling  
 Helping Relationship Inventory (HRI) and the Pilot and Revised  
 Versions of the Medical Helping Relationship Inventory (MHRI)**

	<i>Preliminary Study 1</i>	<i>Preliminary Study 2</i>	<i>Initial Validation Study</i>	<i>Cross-Validation Study</i>
Measurement	HRI	Pilot MHRI	Revised MHRI	Revised MHRI
Class (year)	1986	1987	1988	1989
<i>n</i> subjects	156	154	216	197
<i>n</i> items per scale	10	3	10	10
<i>Response Scales and Coefficient Alpha Estimates</i>				
Understanding	.84	.82	.91	.94
Interpretive	.39	—	—	—
Probing	.58	.50	.64	.68
Focusing	—	—	.62	.66
Supportive/reassuring	.57	.58	.70	.77
Evaluative/judgmental	.53	.44	.70	.69

### INITIAL VALIDATION: STUDIES 3 AND 4

#### SUBJECTS AND PROCEDURE

Two studies were designed to test the reliability (Study 3) and construct validity (Studies 3 and 4) of the final version of the MHRI. The associations between the MHRI scales and students' use of emotional words while responding to patients (convergent and divergent validity) were also tested in Study 3. The sensitivity of the MHRI in measuring changes as a result of an educational intervention designed to enhance students' communication and interviewing skills (decision or construct validity) was examined in Study 4.

Subjects in Study 3 were 216 first-year medical students in the class of 1988. They completed the revised version of the MHRI and another instrument described shortly under the same conditions described for Studies 1 and 2. Subjects in Study 4 were 181 of these students who completed the MHRI both before and after training in medical communication and interviewing skills. Training consisted of 6 hours of didactic presentations, including videotaped examples of communica-

tion techniques and a live physician-patient interview, plus four half-day small-group sessions which included three interviews with elderly residents of community retirement centers and nursing homes. The educational intervention is described in more detail elsewhere (Wolf et al., 1987).

## INSTRUMENTS

The revised version of the MHRI used in Studies 3 and 4 contained two major revisions of the preliminary MHRI used in the pilot study: a) It was lengthened from 3 to 10 items and b) five response options were rank-ordered for each patient case instead of four. The Understanding, Supportive, and Evaluative scales of the pilot MHRI were retained in the final version. Based on the lower reliability of the Probing scale in the pilot version (.50) and the fact that it became clear to us that this scale confounded two distinct types of probing responses, open- and closed-ended questions, the pilot version of the Probing scale was replaced by two distinct scales, one for probing responses and one for focusing responses. A probing response was defined as before but was limited to more open-ended questions, which allow patients to tell their story in their own words and help define the problem as they see it. A focusing response was defined as an attempt to gather specific information that limits patients to a yes/no or one word/short answer response, and was limited to more closed-ended type questions. It is often used to confirm or deny one particular aspect of the problem, typically as a follow-up question or in the review of systems. Both the HRI and the pilot MHRI confounded these two distinct types of responses. Examples of one of the patient cases and the five response types are provided in Figure 1.

Additionally, in Study 3 only, students were asked to respond briefly in one or two sentences to three patient cases similar to the MHRI cases. Their open-ended responses to these cases were then scored for the *presence* (1 point) or *absence* (0 points) of an emotional word using a predefined emotional word list (Wolf et al., 1987). Scores on these patient cases were summed to create the Medical Communication Index (MCI; Wolf et al., 1987), which was patterned after Carkhuff's (1969) Communication Index. Coefficient alpha reliability



---

Patient: *I thought the pain would be gone today, but, boy, it keeps on hurting.*

Physician Responses:

*Can you tolerate the pain?* (focusing response)

*Tell me more about the pain.* (probing response)

*I'm sorry that it's painful longer than you expected. The pain is a natural response of your body to your injury. It shouldn't last much longer.* (reassuring response)

*You sound upset because you are still in pain.* (understanding response)

*I think the medication you are on is most effective and appropriate for treating your problem. Try relaxing—it will make you more comfortable.* (judgmental response)

---

**Figure 1: Example of a Medical Helping Relationship Inventory Patient Case and the Five Communication Response Types That Students Were Asked to Rate From Most to Least Preferred**

for this index was .69. The use of an emotional word is a necessary but not sufficient prerequisite for an empathic or understanding type of response. That is, an empathic response by definition is one where an individual responds to both the content of what the patient says and how the patient feels about it, thus necessitating the use of an emotional word in the response. Because the MHRI Understanding scale is most similar in purpose to the MCI, while the purposes of the other MHRI scales are quite distinct, this index should be positively associated with the MHRI Understanding scale (convergent validity) and negatively or unrelated to the other MHRI scales (discriminant validity).

## RESULTS AND DISCUSSION

Reliability estimates for all the scales on the final version of the MHRI exceeded those for the HRI and pilot MHRI and are summarized in Table 2 under the column labeled "Validation Sample." These coefficients ranged between .62 (Focusing) and .91 (Understanding). The inclusion of additional items and both Probing and Focusing scales on the final version contributed to enhancing the internal consistency of the scales. Validity coefficients between the MHRI scales and the index of emotional responses used are summarized in Table 3. The positive association between understanding and this index (.48) is evidence of convergent validity because the MHRI Understanding scale is most similar in purpose to the MCI. The negative associations between the MCI and the other MHRI scales

**TABLE 3**  
**Pearson Product-Moment Correlations (Validity Coefficients)**  
**Between the Medical Helping Relationship Inventory (MHRI)**  
**Scales and an Index of Emotional Responses Used**

<i>MHRI Response Scale</i>	<i>Validation Sample</i> (n = 216)	<i>Cross-Validation Sample</i> (n = 197)
Probing	-.22**	-.31***
Judgmental	-.24***	-.43***
Understanding	.48***	.69***
Reassuring	-.16*	-.35***
Focusing	-.32***	-.34***

\* $p < .02$ ; \*\* $p < .01$ ; \*\*\* $p < .001$ , two-tailed tests.

provide evidence of discriminant validity, as the purposes of these MHRI scales are quite distinct from that of the MCI. Quite clearly, they are measuring different constructs than the MCI because of these negative associations.

Results of Study 4 are summarized in Table 4 and also lend support to the construct validity of the MHRI. On average, the most preferred response types before training were probing ( $M = 20.21$ ) and supportive (renamed reassuring;  $M = 26.30$ ); the least preferred were evaluative (renamed judgmental;  $M = 37.96$ ) and understanding ( $M = 35.87$ ). After training, understanding ( $M = 18.35$ ) and probing ( $M = 22.05$ ) responses were the most preferred, while judgmental ( $M = 43.50$ ) and reassuring ( $M = 33.55$ ) responses were the least preferred. Understanding types of responses are thought to be reflective of good patient-oriented, sensitive, and humanistic communication (American Board of Internal Medicine, 1983; Carkhuff, 1969; DiMatteo, 1979; Gorlin & Zucker, 1983; Ingelfinger, 1980; Jones, 1973; Nardone et al., 1980), and students exhibited large gains on this scale as a result of training designed to enhance this skill. These gains amounted to almost 2 standard deviation units (1.90), indicating that the average student (i.e., 50th percentile) after training received a score on the MHRI Understanding scale equivalent to the 97.1th percentile ( $U_3$ ) before training. Effect sizes (Cohen, 1988; Wolf, 1986) may be interpreted in general as small ( $d = .2$ ), medium ( $d = .5$ ), and large ( $d = .8$ ).

**TABLE 4**  
**Results of Paired *t*-Tests and Effect Sizes (*d*) for Student Performance on the Medical Helping Relationship Inventory (MHRI) Response Scales Before (Baseline) and After (Follow-Up) an Educational Intervention for an Initial Validation Sample (*n* = 181) and Cross-Validation (Replication) Sample (*n* = 168)**

<i>MHRI</i> Response Scale	Baseline Mean	Follow-Up Mean	Within SD	<i>t</i>	<i>d</i>	<i>U<sub>J</sub></i> (%)
<b>Probing</b>						
Initial validation sample	20.21	22.05	4.78	-5.18	0.38	64.8
Cross-validation sample	21.08	22.65	4.82	-4.19	0.33	62.9
<b>Judgmental</b>						
Initial validation sample	37.96	43.50	4.97	-15.01	1.11	86.4
Cross-validation sample	38.47	42.59	4.86	-10.97	0.85	80.5
<b>Understanding</b>						
Initial validation sample	35.87	18.35	9.21	25.59	1.90	97.1
Cross-validation sample	33.83	19.59	11.26	16.32	1.26	89.6
<b>Reassuring</b>						
Initial validation sample	26.30	33.55	5.63	-17.31	1.29	90.2
Cross-validation sample	26.40	31.37	6.30	-10.19	0.79	78.5
<b>Focusing</b>						
Initial validation sample	29.64	32.54	4.69	-8.31	0.62	73.2
Cross-validation sample	30.28	33.81	4.50	-10.31	0.78	78.2

NOTE: All differences between baseline and follow-up means are statistically significant, *p* < .001, two-tailed tests.

These results indicated that the MHRI scales exhibited the sensitivity necessary to evaluate changes as a result of an educational intervention designed to increase students' communication preferences for understanding responses, decrease their preferences for judgmental and (false) reassuring responses (all large effects), while not diminishing to any great extent their preferences for probing (small effect) and focusing (medium effect) responses.

### CROSS-VALIDATION STUDY AND CONCLUSIONS

A cross-validation study was designed the following year in an effort to replicate the reliability and validity results obtained in the initial validation studies reported earlier here. First-year medical students in the class of 1989 participated in this replication under the same conditions described for the validation studies and are referred to as the cross-validation sample in the tables. Sample sizes ranged between 168 and 197 students for the various replications. Results for the cross-validation sample were comparable for those previously obtained for the validation sample in terms of reliability (Table 2), validity coefficients (Table 3), and utility in evaluating change resulting from intervention (Table 4). An important next step would be to validate performance on the MHRI with actual performance in interviewing real patients.

In conclusion, evidence is presented that suggests a new instrument, the Medical Helping Relationship Inventory, can reliably and validly be used in the evaluation of medical communication skills by assessing preferences for understanding, probing, focusing, reassuring, and judgmental responses to patients. We believe the results of our initial and cross-validation studies warrant experimentation with the MHRI in other settings (e.g., other medical or professional schools) and with other groups of health professionals (e.g., house officers, community physicians, nurses, social workers, and others). We are currently working on the development of additional scales targeted for communication with specific patient populations (e.g., elderly patients; Robins & Wolf, 1989).

## REFERENCES

- American Board of Internal Medicine. (1979). Clinical competence in internal medicine. *Annals of Internal Medicine*, *90*, 402-411.
- American Board of Internal Medicine. (1983). Evaluation of humanistic qualities in the internist. *Annals of Internal Medicine*, *99*, 720-724.
- Carkhuff, R. R. (1969). The prediction of the effects of teacher-counselor education: The development of communication and discrimination selection indexes. *Counselor Education and Supervision*, *8*, 265-272.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*, 292-334.
- DiMatteo, M. R. (1979). A social-psychological analysis of physician-patient rapport: Toward a science of the art of medicine. *Journal of Social Issues*, *35*, 12-33.
- Engler, C. M., Saltzman, G. A., Walker, M. L., & Wolf, F. M. (1981). Medical student acquisition and retention of communication and interviewing skills. *Journal of Medical Education*, *56*, 572-579.
- Gorlin, R., & Zucker, H. D. (1983). Physicians' reactions to patients: A key to teaching humanistic medicine. *New England Journal of Medicine*, *308*, 1059-1063.
- Harris, I. B., Eckert, E. D., Petzel, R. A., & Westermeyer, J. (1984). Evaluation of a medical interviewing course: Use of the helping relationship inventory. *Evaluation & the Health Professions*, *7*, 95-111.
- Haug, M. R., & Lavin, B. (1979). Public challenge of physician authority. *Medical Care*, *17*, 844-858.
- Ingelfinger, F. J. (1980). Arrogance. *New England Journal of Medicine*, *303*, 1507-1511.
- Jarski, R. W., Gjerde, C. L., Bratton, B. D., Brown, D. D., & Matthes, S. S. (1985). A comparison of four empathy instruments in simulated patient-medical student interactions. *Journal of Medical Education*, *60*, 545-551.
- Jones, J. E. (1973). Helping relationship inventory. In J. E. Jones & J. W. Pfeiffer (Eds.), *Annual handbook for group facilitators 1973* (pp. 53-70). San Diego: University Associates.
- Kahn, G. S., Cohen, B., & Jason, H. (1979). The teaching of interpersonal skills in U.S. medical schools. *Journal of Medical Education*, *54*, 29-35.
- Lundberg, G. D. (1985). Medicine—A profession in trouble? *Journal of the American Medical Association*, *253*, 2879-2880.
- McPherson, C., Sachs, L. A., Knopp, W., & Wolf, F. (1984). The doctor-patient relationship: Systematic training in effective communication skills. *Journal of Psychiatric Education*, *8*, 87-92.
- Mechanic, D. (1985). Public perceptions of medicine. *New England Journal of Medicine*, *312*, 181-183.
- Muller, S. (Ed). (1984). Physicians for the twenty-first century: Report of the project panel on the general professional education of the physician and college preparation in medicine [special issue]. *Journal of Medical Education*, *11*(2).
- Nardone, D. A., Reuler, J. B., & Girard, D. E. (1980). Teaching history-taking: Where are we? *Yale Journal of Biological Medicine*, *53*, 233-250.
- Robins, L. S., & Wolf, F. M. (1989). The effect of training on medical students' responses to geriatric patient concerns: Results of a linguistic analysis. *Gerontologist*, *29*, 341-344.

- Stanley, J. C., & Hopkins, K. D. (1972). *Educational and psychological measurement and evaluation*. Englewood Cliffs, NJ: Prentice-Hall.
- Wolf, F. M. (1986). *Meta-analysis: Quantitative methods for research synthesis* (Sage University Paper series on Quantitative Applications in the Social Sciences, 07-059). Beverly Hills, CA: Sage.
- Wolf, F. M., Savickas, M. L., Saltzman, G. A. & Walker, M. L. (1984). Meta-analytic evaluation of an interpersonal skills curriculum for medical students: Synthesizing evidence over successive occasions. *Journal of Counseling Psychologists, 31*, 253-257.
- Wolf, F. M., Woolliscroft, J. O., Calhoun, J. G., & Boxer, G. J. (1987). A controlled experiment in teaching medical students to address patients' emotional concerns. *Journal of Medical Education, 62*, 25-34.