



National Library
of Canada

Bibliothèque nationale
du Canada

Canadian Theses Service Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-315-54949-1

**Interpersonal Skills Competency of
Graduating Medical Students**

by
Gail K. Schnabl

A thesis
presented to the University of Manitoba
in partial fulfillment of the requirements
for the degree of Master of Education in the
Faculty of Education

Winnipeg , Manitoba, 1989

(c) Gail Schnabl, 1989

INTERPERSONAL SKILLS COMPETENCY OF GRADUATING MEDICAL STUDENTS

BY

GAIL K. SCHNABL

A thesis submitted to the Faculty of Graduate Studies of
the University of Manitoba in partial fulfillment of the requirements
of the degree of

MASTER OF EDUCATION

© 1989

Permission has been granted to the LIBRARY OF THE UNIVERSITY OF MANITOBA to lend or sell copies of this thesis, to the NATIONAL LIBRARY OF CANADA to microfilm this thesis and to lend or sell copies of the film, and UNIVERSITY MICROFILMS to publish an abstract of this thesis.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

ABSTRACT

The study reported here assessed the interpersonal skills of the 1988 graduating class of the Faculty of Medicine of the University of Manitoba. This assessment comprised one component of a larger comprehensive clinical examination being given to these students in December, 1987. In view of the importance of interpersonal skills to physicians' practice, it is essential that medical schools evaluate the competency of their graduands in this area.

The format of the examination consisted of 20 performance-based assessments or stations, 19 of which were live simulations. Live simulations utilize individuals, who are trained to present, in a standardized way, the physical signs and symptoms of a problem for which typically a patient might seek medical care or advice. The simulated patients were interviewed and/or physically examined by the medical students. Following each encounter, the simulated patient assessed the medical student's performance in a number of areas including interpersonal skills using checklists.

Utilizing the assessments by the simulated patients, this study evaluated the interpersonal skill level of the students, and investigated whether or not competency in interpersonal skills was correlated with other medical competencies, whether it varied across medical problems and

finally, how accurately the students themselves could assess their own interpersonal skills. The investigation found generally that the students' interpersonal skills were assessed favourably by the simulated patients with a few students identified as having deficiencies. Secondly, the study found that interpersonal skill competency correlated moderately with skills in data gathering. The investigation also found that interpersonal skill competency varied significantly ($p < .05$) across the medical problems greater than was accounted for by rater variability. Finally, medical students with higher levels of interpersonal skills were better able to predict how a simulator would assess them. Implications for the teaching of interpersonal skills and research recommendations are provided.

ACKNOWLEDGEMENTS

I would like to extend my gratitude to several people who have supported me during the completion of my thesis. Thank you to my advisor, Dr. Clare Pangman, for his on-going encouragement and guidance, and to my committee members, Dr. Daniel Klass, Dr. Thomas Hassard and Dr. Hal May for their direction and assistance. It was the inspiration and support of Dr. Klass, that led to my involvement in the Clinical Comprehensive Examination. I also wish to acknowledge the special assistance of Dr. Hassard through whose guidance the data analysis became a reality.

As well, I would like to thank Dr. Murray Kopelow, Vera Siemans and others involved in the administration of the Clinical Comprehensive Examination for their support and assistance in collecting the data.

I would like to thank the Office of Educational Development and the Department of Family Medicine of the Faculty of Medicine for their financial support with the costs of data analysis.

I would like to extend my appreciation to all the medical students and simulators who participated in the Clinical Comprehensive Examination and whose data contributed to this study.

I would like to acknowledge the understanding and continuous support of my husband, Johann, and my son, Peter. I would like especially to thank Johann for introducing me to "the computer" and for his on-going consultation which made completion of this thesis so much easier.

TABLE OF CONTENTS

i	Abstract
iii	Acknowledgements
v	Table of Contents
vii	List of Tables

CHAPTER

I	INTRODUCTION	1
	General Background	1
	Statement of the Problem	5
	Research Questions	7
	Definition of Variables	9
II	REVIEW OF THE LITERATURE	13
	Components of Interpersonal Skills	13
	Assessment of Interpersonal Skills	16
	Physician Competence in Interpersonal Skills	19
III	METHOD	23
	Subjects	23
	Instrument Used	23
	Procedure	26
	Delimitations.....	28
	Data Analysis	28
IV	RESULTS	34
	Presentation of Findings	34
	Research Question 1	35

	Research Question 2	39
	Research Question 3	43
	Research Question 4	48
	Supplementary Findings	50
	Discussion	57
	Research Question 1	57
	Research Question 2	61
	Research Question 3	64
	Research Question 4	67
	Supplementary Findings	69
V	CONCLUSIONS, LIMITATIONS AND	
	RECOMMENDATIONS	71
	Conclusions	74
	Limitations	76
	Recommendations	78
	Research Recommendations	80
	References	82
	List of Appendices	
	A Interpersonal Skills Rating	
	Scale.....	92
	B Communication and Professional	
	Service Checklist	94
	C Letter to Medical Students	95
	D Correlation Coefficients between	
	Interpersonal Skills Rating	
	Scale and Communication and	
	Professional Service Checklist	96

E	Instructions to Students for Predicting Simulator's Assessment.....	97
F	Scatterplots of Original and Modified Scores, for Inter- personal Skills, Data Collection and Management.....	98

List of Tables

Table 1	Interpersonal Skill Scores U. of M. Students	36
Table 2	Comparisons of Original and Modified Competency Scores Using Correlational Coefficients	38
Table 3	Relationships between IPS Scores (Original and Modified) & Other Competencies, Using Correlational Coefficients	41
Table 4	Coefficients of Determination (r^2) between IPS & Other Competencies Using Original and Modified Scores ...	43
Table 5	Random Effects Analysis of Variance Summary Table.....	45
Table 6	Mean IPS Scores by Station, Grouped by Similarity	47
Table 7	Eigenvalue Summary IPS Rating Scale	51
Table 8	Rotated Factor Loadings IPS Rating Scale	52
Table 9	Eigenvalue Summary IPS Rating Scale & Communication and Professional Service Checklist	54
Table 10	Rotated Factor Loadings IPS Rating Scale and Communication and Professional Service Checklist	56

CHAPTER I

INTRODUCTION

General Background

Interpersonal skills of physicians are an important component of an effective doctor-patient relationship. The interpersonal skills required for interviews with patients need to be delineated from the interpersonal skills physicians may require for other aspects of their role. Specifically, the interpersonal skills required for interviewing patients include sensitivity to patients' needs, the ability to establish rapport, and other skills which might be described in a general way as skills of communication. These latter communication skills involve techniques of gathering information from patients face to face and, in turn providing explanations to patients about diagnoses and treatments. On the other hand, while relationships with medical colleagues and other health professionals may require some of these same skills, such interactions may also require additional skills such as greater directness, competence with medical jargon and a matter-of-factness. However, it is the former set of interpersonal skills, those relevant to interactions with patients, which are the focus of this study.

Interpersonal skills, such as providing warmth in the relationship, using understandable language and giving clear explanations, have been shown to influence patient satisfaction and compliance (Cousins,1985; DiMatteo,1979; Fisher,1971; Francis, Korsch & Morris,1969; Korsch, Gozzi & Francis,1968; Sanson-Fisher & Maguire, 1980; Sideris, Tsouna-Hadjis, Toumanidis, Vardas & Mouloupoulos,1986; Stiles, Putman, Wolf & James,1979; Stone,1979). Furthermore, a lack of such skills can often be a factor in law suits against physicians (Messenger, 1987; Vaccarino, 1977) where, for example, physicians may not explain the predicted outcome of treatment either adequately or in understandable language and patients feel improperly cared for. It appears evident from the literature that interpersonal skills are relevant to most doctor-patient interactions with the possible exception of those requiring an emergency response. Given the essential importance of interpersonal skills, it seems crucial that graduates of medical schools be evaluated as competent, not only in all the medically related skills, but in interpersonal skills as well.

Until recent years, the active development of interpersonal skills in medical students was limited (Kahn, Cohen & Jason,1979). There is still considerable variation among medical schools as to the amount of curriculum and faculty time that is devoted to the development of such skills (Carroll & Monroe,1980). Most courses on

interpersonal skills are provided in the pre-clinical years of medical training; that is, in first or second year. The subsequent experience of the medical student as a clinical clerk may not reinforce this early learning. Both the pressure to integrate a constantly expanding body of knowledge and the high profile given to technological medicine, may challenge the students' efforts to place adequate emphasis on the skills of establishing rapport and communicating with patients. Consequently, at graduation, medical students are often unable to demonstrate appropriate competency in interpersonal skills with patients (Knox, Alexander, Morrison & Bennett, 1979; Maguire & Rutter, 1976).

Traditionally, commencement examinations for medical students have been multiple choice in character. This type of question does not assess interpersonal skills. On the other hand, on-going assessments by supervising physicians of students' interactions with patients lack comparability as patients are not standardized and evaluators may have different performance expectations (Harden & Gleeson, 1979; Harper, Roy, Norman, Rand & Feightner, 1983).

Recently efforts have been directed to standardized performance-based assessments of medical students (Abrahamson, 1985; Williams et al., 1987). One variation utilizes individuals trained to present a specific patient problem. A description of symptoms, personal information

and physical signs are carefully taught to these individuals so that they can consistently present this information in response to questioning. These individuals are then considered simulated or standardized patients. The medical student interviews and/or carries out a physical examination on these simulated patients and then decides on a course of action. Simulated patients have been found to provide a valid representation of a patient's problem (Norman, Tugwell & Feightner, 1982; Sanson-Fisher & Poole, 1980). In the latter study, students were unable to discriminate between simulated and genuine patients. Provided that an adequate number of examples of a medical student's performance are assessed, this approach is reliable. With fewer than 15-20 examples of a student's performance, performance ability cannot be generalized (Stillman et al., 1986).

Some of the advantages of using simulated patients in the assessment procedures should be indicated. The simulated patient can provide valuable feedback information on the student's performance. In particular, the simulated patient can assess the student's interpersonal skills. In addition to determining the general level of competency of each student as a medical practitioner, students can be compared and their competency in interpersonal skills can be compared to their competency in other skills, such as data collection, diagnosis and management.

Through the use of simulated patients it may also be possible to determine whether or not medical students' skills in interpersonal relationships with patients vary as a function of the patient problem or if this ability remains fairly constant and unaffected by the nature of the patient's problem. If medical students are perceived by simulated patients as being less effective, interpersonally, in dealing with certain medical problems either as individuals or in a group, such observations, if valid and reliable, would have implications for a medical student's education and possibly for the curriculum of the medical school concerned. Alternately, if interpersonal skill competency is found to be fairly constant across problems, it would then be necessary to assess only a limited sample of a student's interviewing performance to determine competency in interpersonal skills.

A final advantage to using simulated patients is as a means for self-assessment. Students can evaluate their own interpersonal skills and such evaluations can be compared to those completed by the simulated patients.

Statement of the Problem

The University of Manitoba and Southern Illinois University have cooperated over the last two years to mount a jointly drafted comprehensive clinical examination for the fourth year graduating medical students at each university.

The first cooperatively drafted examination was administered for the 1987 graduating classes of both universities (Williams et al., 1987). This examination involved a number of stations which utilized simulated patients and carried out a limited assessment of the medical students' interpersonal skills. The interpersonal skills were assessed on less than 25 % of the stations on that first examination. In addition, the interpersonal skills checklist consisted of only 6-8 items and was not consistent from one station to another. It was agreed by the organizers of the examination that further assessment of the students' competency in interpersonal skills should take place in the next examination.

The investigation reported here was exploratory and designed : (1) to assess the interpersonal skill competency of a group of graduating medical students, (2) to compare the interpersonal skill competency to other competencies of medical practice, (3) to compare the interpersonal skill competency across medical problems and (4) to determine how accurately medical students could predict simulated patients' assessment of their interpersonal skills. The interpersonal skills were assessed across all 19 simulations involving standardized patients. This number of stations provides more than an adequate sampling of student behaviour from which to generalize about student ability (Stillman et al., 1986).

Research Questions

Specifically, the study focused on the following research questions.

1. What is the interpersonal skill (IPS) competency level of the graduating medical students at the University of Manitoba as a group?

Effective interpersonal skills are important for physicians, and as such, it is essential that graduating medical students be assessed in this area.

In that such an assessment of interpersonal skills has not been completed previously, determining what is competent or effective performance by comparing the results against an external standard or established norms poses a problem. This study may, in fact, contribute to the establishment of norms for this area of medical skill. The issue of determining interpersonal skill competency will be examined further when the results are discussed.

2. What is the relationship, if any, between the level of IPS competency and other competencies, specifically: (1) data collection (2) diagnosis (3) management and (4) overall non-interpersonal skill competency?

It is of educational relevance to determine the extent to which competency in interpersonal skills is associated with competency in other areas. For instance, it can be

useful to discover if effective interpersonal skills enable physicians to obtain more useful information from patients through the interview or physical examination.

3. Does the level of interpersonal skill competency vary significantly as a function of the particular patient problem that the simulated patient presents?

If interpersonal skills vary as a function of the patient problem and there are noted deficiencies, then there are implications for medical education for individual students and for the medical school as a whole. If there is no significant variation across problems, assessment of interpersonal skills in future could be carried out using fewer examples of a student's performance.

4. What is the predictive accuracy of the medical students at assessing their own interpersonal skill competency?

This predictive accuracy is one measure of self-assessment skills. Self-assessment is gaining more credibility and importance in the professional development of both educator and professional. Graduating medical students will undoubtedly use self-assessment in their future practices. The ability to perceive possible problems in interactions with patients is the first step to correcting such problems and could lead to improved patient care.

Having outlined the research questions, it is now necessary to identify and define the relevant variables which will be encountered in the study.

Definition of Variables

Nine variables require definition : (1) interpersonal skill competency, (2) simulated patients, (3) graduating medical students, (4) patient problems, (5) data collection competency, (6) diagnosis competency, (7) management competency, (8) overall competency in non-interpersonal skills and (9) predictive accuracy.

1. Interpersonal Skill Competency

For purposes of this study, "interpersonal skill competency" is defined as the student's score on the rating scale completed by the simulated patient following an encounter with that medical student. The simulator based the assessment on his or her perception of the student's performance in that encounter. The rating scale assessed two main areas of interpersonal skills (1) empathy and (2) communication. The relevance of these two components of interpersonal skills for physicians' performance is discussed in the review of the literature.

Selected items of the empathy scale of the revised Barrett-Lennard Relationship Inventory were used to measure empathy (Barrett-Lennard, 1978). The component of

communication was assessed by items directly soliciting perceptions by the simulators of the medical students' communication. The items are those commonly found in scales used by other researchers.

2. Simulated Patients

The "simulated patients" are individuals who were trained to portray in a standardized way symptoms and signs of a patient problem. At least two simulators were trained for each problem. This apparent duplication of simulators was found to be necessary for two reasons. In order to analyze the data for any difference in IPS due to case or station differences (research question 3), it was necessary to have at least two raters for each station so that differences due to variability in raters could be calculated. Secondly, two simulators were required to reduce possible fatigue given that approximately 100 repetitions were required of each simulation.

3. Graduating Medical Students

The "graduating medical students" were the 92 students in the 4th year class of the University of Manitoba Faculty of Medicine, who graduated in May, 1988.

4. Patient Problems

The "patient problems" were those the simulators were trained to present. Nineteen such problems, based on

problems which actual patients have presented to physicians, were used in the study.

5. Data Collection Competency

The authors of each station in the examination designated specific items as measuring competency in data collection. Therefore, for this study "data collection competency" is defined as the score given to each student on items designated as data collection on the examination. A large part of the data collection score is comprised of assessments by the simulator of items of history asked or physical examinations performed by the student. In addition, in some stations paper and pencil items completed by the student are also part of this competency.

6. Diagnosis Competency

The "diagnosis competency" is defined as the score given to each student on the items designated as measuring diagnosis on the examination. The diagnosis score is derived from paper and pencil items completed by the student.

7. Management Competency

The "management competency" is defined as the score given to each student on items designated as measuring management on the examination. The management score is derived from either assessments by the simulators on some

stations or from paper and pencil items completed by the students on other stations.

8. Overall Non-Interpersonal Skill Competency

"Overall non-interpersonal skill competency" (ON-IPS) is the score the student received when all the other components of the examination except those pertaining to interpersonal skills were combined.

9. Predictive Accuracy

"Predictive accuracy" is defined as the difference in scores between a student's prediction of how a simulated patient would assess him/her and the actual simulator's assessment of that student.

CHAPTER II

REVIEW OF THE LITERATURE

The review of the literature is organized and presented in three areas: (1)the components of interpersonal skills, (2)the assessment of interpersonal skills and finally (3) physician competence in interpersonal skills.

Components of Interpersonal Skills

The interpersonal skills addressed in this study need to be identified. The first of two main components, which emerge from a review of the literature, involves the quality of empathy between the patient and the doctor. The second includes certain aspects of verbal communication. Each component is described in more detail beginning with empathy.

Empathy has been defined as "the ability to apprehend another person's feelings and to psychologically assume another person's role without enacting that role" (English & English,1958, p.178). Empathic understanding allows the physician to be sensitive to the patient's needs and allows for a sense of rapport to develop between them. This quality of empathy has been found to be important to the establishment of a relationship between client and therapist in psychotherapy and counselling (Barrett-Lennard,1962; Carkhuff & Berenson,1967; Rogers,1957). The ability to

establish rapport with patients is cited in the literature as being an essential quality for physicians (Francis et al., 1969; Poole & Sanson-Fisher, 1979; Ward & Stein, 1975) and one which patients expect their physicians to possess (Mace, 1971). DiMatteo (1979) reported that patient satisfaction was correlated with the physician's ability to read the patient's emotions and with the ability to communicate emotions non-verbally to the patient. Of special significance to this investigation is the Sanson-Fisher & Poole (1980) study where no significant differences were found between the levels of empathy exhibited by students in interactions with genuine patients and empathy levels with simulated patients. In that the present study utilized simulated patients for assessing students, one can reasonably assume students demonstrated as much empathy towards them as they would have with genuine patients.

The second component of communication comprises the use of appropriate and understandable language. It includes the ability, after adequate information has been gathered, to inform the patient appropriately of the diagnosis and to explain clearly the treatment or other management. The importance and value of these skills in medical encounters have been raised frequently (Francis et al., 1969; Korsch et al., 1968; Stewart, 1984; Stiles et al., 1979; Stone, 1979; Waitzkin, 1984). Korsch and her colleagues found that dissatisfaction with medical encounters was frequently

related to a lack of information received by mothers about their child's illness and to the use of medical jargon. Stiles et al. (1979) reported that one of the main factors affecting patient satisfaction was the physician giving information to the patient and providing the patient with an opportunity to ask questions about this. Stewart found higher compliance and satisfaction by patients in interviews which she describes as patient-centred. Patient-centred interviews are characterized by patients having an opportunity to express their feelings and to give suggestions to the physician and by the physician seeking such opinions out and providing support. Patient centredness has features of both empathy and effective communication.

In another investigation (Hassard, Kopelow, Schnabl & Klass, 1988) two factors, sensitivity and participation, were reported to be associated with simulated patients' satisfaction with care received from medical students. The sensitivity factor as analyzed has many features of empathy. The participation factor is described as involvement of the patient through explanations of the medical problem and in the planning of treatment, in other words the communication component.

A brief review of the literature indicated above seems to suggest that it is important for practicing physicians to demonstrate both empathy and effective communication in their interactions with patients. Because of the importance

of empathy and verbal communication these two components formed the basis of the assessment of the medical students' interpersonal skills in this study.

A description of how interpersonal skills have been assessed by other investigators is provided next.

Assessment of Interpersonal Skills

Interpersonal communication skills of student physicians and other health professionals have been assessed using a number of different approaches. Often these approaches utilize trained observers or interpersonal communication experts to look for the presence of specific behaviours which are considered to be indicative of appropriate interpersonal skills (Engler, Saltzman, Walker & Wolf, 1981; Irwin & Bamber, 1984; Maguire, Clarke & Jolley, 1977). Medical-student interviews, either live or video-taped, are assessed by the use of rating scales or check-off lists. Often such standardized feedback information is provided to medical students in teaching sessions where evaluation is formative rather than summative in nature. Formative evaluation is evaluation provided periodically during the course of learning whereas summative evaluation refers to a terminal assessment. Assessments conducted by outside observers have been shown to be reliable. However, the results do not always correlate strongly with how positively the patient or simulator

regards the interviewer or how satisfied the simulator feels about the interview (Felletti & Carney, 1984; Holmes, Baker, Torian, Richardson, Glick & Yarmat, 1978).

Assessment by simulated patients of the quality of interpersonal communication skills demonstrated by medical students has been a clearly established practice for a number of years. The use of simulated patients in both teaching and formative assessments has been well documented (Coulehan & Nardini, 1982; Faber, Out & Reepmaker, 1984; Froelich, 1969; Hannay, 1980; Maguire et al., 1977; Simek-Downing & Quirk, 1985; Stillman, Burpeau-Di Gregorio, Nicholson, Sabers & Stillman, 1983).

There has been less frequent use of simulated patients in summative assessments of interpersonal skills. One reason for the infrequent use of standardized patients is the increased faculty and student time required and the additional expense involved in a performance-based format for summative assessment. Another reason could be the common reluctance to move to new and unfamiliar approaches.

Although less frequent than in formative evaluation, the use of simulated patients in the summative assessment of interpersonal skills has been shown to be possible and appropriate. In fact, simulated patients were used in a summative assessment of internal medicine residents in 14 U.S. medical schools and this assessment included a

communication skills component (Stillman et al., 1986). A similar pilot assessment of fourth year medical students has also been reported (Stillman, Regan & Swanson, 1987). However, the pioneer in the field of using live simulations for assessing physician performance is the College of Family Physicians of Canada. The College has assessed affective skills competency through simulations on its certification examination for the past 20 years. Over the years the simulators have been actors and most recently other physicians have simulated the particular patient condition required. In all of the above mentioned examples, the components, defined as communication skills or affective skills, are similar to, although perhaps not identical to, the interpersonal skills which have been delineated for this study.

The performance-based OSCE (Objective Structured Clinical Examination) also can be used to assess interpersonal skills through simulations by physicians or other individuals. Within the format of an OSCE, skills are assessed separately in that the simulations do not attempt to replicate a complete doctor-patient encounter but only a short segment of it. For example, the student might be given 5 minutes to explain a diagnosis to a patient. These examinations are often used summatively as terminal assessments of competence (Harden & Gleeson, 1979).

The review of the literature suggests that while summative assessment of interpersonal skills does not occur on a regular basis for graduating medical students, it is feasible utilizing a performance-based tool and using simulated patients. The present study has confirmed that it is both possible and practical to include an assessment of interpersonal skills within a performance-based summative examination and to train simulated patients to carry out such an assessment.

A description of the interpersonal skill competency of practicing physicians as presented in the literature will be discussed next.

Physician Competence in Interpersonal Skills

Earlier it was mentioned that interpersonal skills are taught in the first or second year of medical training. Some reports claim, however, that what is learned about interpersonal skills is not necessarily retained by the time of graduation (Engler et al., 1981; Kauss, Robbins, Abrass, Bakaitis & Anderson, 1980; Sanson-Fisher & Poole, 1979). Studies of practicing physicians have also shown that they may be significantly deficient in this area (Byrne & Long, 1976; Maguire, Fairbairn & Fletcher, 1986). Furthermore, doctors do not always accurately assess their own interpersonal skills. In fact, it has been shown that they tend both to underestimate patients' needs for information

and to overestimate how much information they provide (Waitzin,1984).

Of some relevance to this investigation, research completed with psychology students showed that highly empathetic individuals were more likely than less empathetic individuals to make accurate predictions of how others perceived them (Harman,1986). The present study examined the question of whether or not medical students were able to predict accurately how a simulated patient would assess them.

Other related research has shown that some physicians may have difficulty in dealing with non-medical issues that arise in interviews with patients (Noren, Frazier, Altman & DeLozier,1980). Frequently, patients bring concerns regarding life situations or emotional problems to physicians. Discussion of these non-medical issues may be relevant to medical treatment. It may well be that the interpersonal skill competency of the physician has a bearing on whether he/she feels comfortable handling such non-medical problems.

Some comparisons have been carried out between interpersonal skill competency of residents and their competency in other aspects of a medical encounter such as history taking, physical examination and diagnostic skills (Stillman et al.,1986). Moderately high positive

correlations were found between communication skills and those of history taking and also physical examination but there was less of a relationship between communication skills and diagnostic skills. Interestingly, other investigators, (Wolraich, Albanese, Reiter-Thayer & Barrett, 1982) found that increasing knowledge about a medical problem was not associated with improved interviewing skills with patients who had that problem.

Stillman (1986) and her fellow researchers compared the communication skills of individual medical residents over three or four different patient problems and found that mean scores varied across cases. The College of Family Physicians of Canada has collected data on affective skills during its annual examination of family medicine residents and practicing physicians for many years. This data suggests that affective skills, as assessed by the College, may vary as a function of the problem content (P. Rainsberry, personal communication, October 9, 1987). It seems both relevant and timely that further research into this area be conducted now with graduating medical students because of its implications for medical education.

This review of the literature confirms that effective interpersonal skills are important in interactions between physicians and patients. Compliance is enhanced and patients are more satisfied when physicians make efforts to understand and communicate with them. Therefore, assessing

interpersonal skill competency in graduating medical students is crucial. The literature also supports the use of simulated or standardized patients in carrying out this assessment. The next chapter will describe how the assessment of interpersonal skills was carried out in this investigation.

CHAPTER III

METHOD

Subjects

The subjects for this study were the 1988 graduating class of the Faculty of Medicine at the University of Manitoba. These students undertook a comprehensive performance-based examination in December, 1987. The examination took place in an former ward of the Children's Centre, Health Sciences Centre which was set up to simulate clinical rooms for seeing patients. The assessment of interpersonal skills was one component of the examination, which was comprised of 19 live and one paper simulation. The entire class of 92 students was assessed at these stations over a 2 week period. In addition to the 92 students, 7 graduates of non-Canadian medical schools undertook this examination as part of a process of obtaining Canadian credentials. They are also included as subjects for some aspects of the investigation.

Instrument Used

The instrument used to assess interpersonal skills was a 14 item rating scale completed by the simulated patients. Six items were designed to assess empathy; six additional items formed the communication component; and one item each assessed the simulated patient's perception of thoroughness

and satisfaction. The six point rating scale ranged from +3 to -3. Individual items were positively or negatively associated with effective interpersonal skills. A copy of the scale is included in Appendix A.

The empathy items were derived from the revised Barrett-Lennard Relationship Inventory (Barrett-Lennard, 1978). Although several different instruments have been used to assess empathy, this particular instrument was chosen for several reasons. Its reliability and validity are both considered acceptable. Barrett-Lennard (1962) originally reported split-half reliabilities of .86 for the empathy sub-scale and subsequent studies have reported split-half and test-retest reliabilities in the same range (Harman, 1986; Ponterotto & Furlong, 1985). The scale also appears to have reasonably good predictive validity as measures of empathy on this scale have been moderately related to outcome (Barrett-Lennard, 1981). The Barrett-Lennard Relationship Inventory has also been used in the assessment of counselling relationships (Barrett-Lennard, 1962; Claiborn, Crawford & Hackman, 1983; Meen, 1986; Mills & Zytowski, 1967) and in the evaluation of single medical encounters (Jarski, Gjerde, Bratton, Brown & Matthes, 1985; Malpiede, Leff, Wilson & Moore, 1982). It lends itself to assessment by simulated patients because it was designed to be used by the interviewee rather than by outside observers.

The verbs in the scale items in the Relationship Inventory are written in the present tense. For this study the verb tense was changed to the past to reflect the reality of assessing an encounter which has been completed as opposed to a relationship which may continue. Such a change assists the understanding of the individual completing the form and is in keeping with what was done in another study of single medical encounters (Jarski et al., 1985).

The Barrett-Lennard Relationship Inventory has 16 items to measure empathy. The original plan had been to use all 16 items for this investigation. However, due to training time restrictions and the limited time available to simulators for assessing students between encounters (five minutes), it was decided to modify the scale and to limit the empathy items to six. The specific six items were chosen by the investigator as being those most relevant to single medical encounters. This does raise the question as to how completely empathy was measured but the decision to compromise seemed to be the only practical approach to deal with the realities of the situation.

Six items provided for the simulator's assessment of communication skills. It was not felt necessary to use a complete established instrument to assess the communication component as the skills are very specific ones and are being assessed directly, unlike in the case of empathy which by

its very nature must be measured indirectly. The actual items included in the rating scale are those found commonly in scales used to measure patients' perceptions of physicians' verbal communication in interviews. A copy of the Interpersonal Skills Rating Scale is included in Appendix A.

Procedure

The simulated patients were trained for their roles by the investigator and clinical experts designated for each station according to a method originally developed by Barrows (1971). The simulators were paid at a rate of \$10.00 per hour for time spent in training and actual patient simulation. The cost of the simulators was born by the Faculty of Medicine as part of the cost of the whole examination.

The simulators were also oriented by the investigator to the rating scale that they were to use to assess the students' interpersonal skills. The items on the scale were explained and the general protocol for assessing the students was described. It had been planned to give the simulated patients an opportunity for at least two practice sessions assessing physicians during a rehearsal of their station. However, this occurred in less than one-half of the stations due to time limitations. Any limitations this may have placed on the results will be discussed in Chapter V.

No attempt was made to achieve inter-rater reliability. Any significant differences in simulators' ratings were instead handled in the analyses of the data and will be described in the section on data analysis.

In addition, the simulators were trained to complete checklists to assess whether or not a student had gathered certain information or had carried out certain physical examination procedures. The simulators completed the rating scale and appropriate checklists in the few minutes following the encounter with each student.

After the last encounter with a simulated patient, each student was asked to complete the interpersonal skill rating form and to rate themselves as they would predict the simulator in the last previous encounter would have assessed them. The instruction sheet given to the students to complete this task is included in Appendix E. This formed the self-assessment component and provided the information to produce the predictive accuracy score. It was important that the students complete this rating after the last encounter, rather than before any encounter, because a preview of the interpersonal skill rating form would be almost certain to have biased the performance of the medical students. The fact that the students predicted the assessments they received from different simulators was not problematic because the critical issue was not how well they

were assessed by the simulator but how closely they were able to predict the simulator's assessment.

Delimitations

This study examined single encounters between a student doctor and a simulated patient and did not assess a relationship over time. Patients' perceptions of physician skill may change as the relationship develops over several interactions.

This investigation did not consider any measures of outcome of the doctor-patient encounter. It was not possible, for instance, to determine if highly rated interpersonal skills were related to effectiveness as measured by compliance with medical regimens.

Finally, no attempt was made to assess the accuracy of the explanation given by the student to the patient in communicating a diagnosis or treatment plan. As a result, the student may have been rated highly but have given incorrect information to the simulator.

Data Analysis

The data were coded and placed in a data set for analysis. The raw student scores were modified to convert them to a manageable form. This was necessary because the marking sheet contained negative numbers, that is, the possible scores were -3,-2,-1, +1,+2,+3. Thus, it was

possible for a student to receive a negative total IPS score. Therefore, all item scores were modified by adding +3 to the student's score making the possible range of scores for each item 0 to 6. The total maximum for IPS then became 84. The range for each of the components of IPS became automatically modified. Empathy and communication, each with 6 items, had a maximum score of 36 each. Both thoroughness and satisfaction, with one item each, have a maximum score of 6. Analyses of variance, correlations and factor analyses were performed on a personal computer utilizing the Number Cruncher Statistical System (Hintze, 1986).

The data were analyzed initially on an individual station basis and then they were combined across stations. Means and standard deviations were obtained for the competencies for each station. This included the IPS score and its components (Empathy, Communication, Thoroughness and Satisfaction) and other competencies such as Data Collection, Diagnosis and Management. Analyses of variance were carried out to test for any difference in the means of the raters for each station for each of the components. A level of significance of .05 was used for testing for rater difference. Using a higher level of significance such as .01 would have increased the likelihood of failing to find a difference in raters when, in fact, such a difference did exist (Type II error). Therefore, the .05 level was considered more appropriate in this instance than a level of

.01 or higher. Where there were significant differences in rater mean scores, the individual student scores were adjusted. This was done by adding or subtracting from the student's scores an amount equal to the difference between their rater's mean and the station mean. The result was two sets of scores for each student: an original score and an adjusted score. Correlational analyses (Pearson's and Spearman's) were carried out on these sets of scores to test their similarity to each other.

Correlational analyses (Pearson's) were carried out on the competencies of IPS, Data Collection, Diagnosis and Management to test the relationship between them. In addition, similar correlational analyses were done using the components of IPS. The purpose was to determine whether or not competence in IPS, or in any of its specific components, was associated with competence in data collection, diagnosis or management.

An Overall Non-Interpersonal Skill (ON-IPS) competency score was computed for each student. This was done by subtracting from the overall competency score the IPS score and, in addition, that portion of the total marks attributed to Communication and Professional Service. The latter two skills were competencies which Southern Illinois University had used in their examination and which the University of Manitoba had also included in its examination for purposes of comparing results. The investigator believes that these

two competencies reflect interpersonal skills, so they were excluded from the ON-IPS score. Correlations were then carried out between the ON-IPS competency scores and IPS scores. The purpose was to determine whether or not competency in IPS was associated with competency in medical skills not involving IPS.

Differences in IPS scores between students, between raters within stations and between stations were tested using a random model of ANOVA to determine if there was some difference between the stations which could not be accounted for by the fact that there were different raters assessing students at every station. A random model was used because the students, raters and stations must be considered samples drawn at random from a very large population, the limit of which is not exactly known (Ferguson, 1971). It was reasoned that, if such a difference was found, a considerable portion of the difference would be due to the content of the problem presented at the station. A significance level of .05 was used to assess these differences. Tukey's critical values were utilized to determine which specific stations differed from each other because multiple pairwise comparisons were involved. Rank-order correlations were then carried out between the age of the simulated patient and the mean IPS score for the station to determine whether or not there was any relationship between IPS competency of the students and the age of the patients. Rank-order correlations were also

done using the mean IPS score and the mean overall competency score for the station to test for any association between them.

A predictive accuracy score for each student was obtained by calculating the difference between the simulator's score for the student and the student's prediction of that score. Group mean scores were calculated for the students who under-rated their IPS scores and for the students who over-rated their scores. (There were no students who predicted with complete accuracy how the simulators would assess them.) The mean scores of the two groups were then compared using ANOVA to determine whether or not there was a difference in interpersonal skills between these two groups. Pearson's product-moment correlations were also carried out between the predictive accuracy scores and: (1) the mean IPS scores and (2) the mean scores of the components of Empathy and Communication. These correlations were done to see if there was a relationship between competence in IPS and the ability to predict how another person has assessed one's interpersonal skills.

Although not part of the originally planned investigation, both the set of individual items of the IPS scale and the combined items of IPS and the Communication and Professional Service Checklist from Southern Illinois University were factor analyzed. A copy of the Communication

and Professional Service Checklist is included in Appendix B. The factor analysis of the IPS scale was done to find out whether or not the scale did indeed measure four separate factors as had been planned. The Southern Illinois University Checklist was then included in the factor analysis to determine if there were additional factors measured by that checklist which were not part of the IPS scale. An eigenvalue value of 1 or greater was used to select factors in both instances as this is a generally accepted criterion for extracting factors from a matrix of data (Child, 1970; Kim & Mueller, 1978). An eigenvalue is a mathematical property of a correlational matrix used in the decomposition of a matrix into its common factors.

CHAPTER IV

RESULTS

Presentation of Findings

The complete examination of 20 stations included 19 simulations involving standardized patients and 1 paper simulation. Only the 19 "live" simulations assessed interpersonal skills. It was necessary to exclude three stations from analyses combining data from all stations. In two cases, not all students were able to complete the stations because the simulator was unavailable. In one other case, interpersonal skills were assessed according to an abbreviated form of only 4 items rather than 14 because of the nature of the encounter between student and patient (over the telephone). The results presented are, therefore, on 16 stations.

Ninety-nine individuals took the examination. Data from four students were excluded from the analyses because they did not complete all 16 stations. Of the 95 individuals who did complete the examination, 7 were graduates of non-Canadian medical schools who were sitting the examination as part of a process of obtaining Canadian credentials and whose results had been combined with that of the University of Manitoba graduating students. Overall mean IPS scores for this group of 7 examinees were compared to the overall mean IPS scores of the 88 University of Manitoba students using

ANOVA and were found to be significantly different ($F = 7.16$; $df = 1,93$; $p = .009$). Therefore, the results presented here will differentiate the scores of the University of Manitoba graduating students from the total group when it is appropriate and possible to do so.

Research Question 1

To answer the question "What is the interpersonal skill level of the graduating medical students as a group?", the mean overall IPS score for the 16 stations was calculated for the group of University of Manitoba students and found to be 65.78 (S.D. 4.35) or 78.31%. The median score is 66.47 with a range of 23.69 from a low of 50.50 to a high of 74.19. These results along with the results of similar calculations for the components of IPS are presented in Table 1.

For the remaining analyses of the data related to this question, scores of both the University of Manitoba and the non-Canadian graduates are combined. This is appropriate because the analyses involve examining the effects of different raters on scores and not the competence of the students in the area of interpersonal skills.

TABLE 1
INTERPERSONAL SKILLS SCORES
UNIVERSITY OF MANITOBA STUDENTS (N = 88)

Competency	Mean (S D)	Median	(Range) (low-high)
Total IPS (84)*	65.78 (4.35)	66.47	(50.50-74.19)
Empathy (36)*	28.19 (1.87)	28.38	(21.06-31.50)
Communication (36)*	28.54 (1.89)	28.69	(22.88-32.31)
Thoroughness (6)*	4.19 (.55)	4.25	(3.00-5.38)
Satisfaction (6)*	4.82 (.41)	4.88	(3.69-5.69)

* Maximum score possible appears in brackets.

One of the major issues in the analysis of the data was that for each station there were at least two raters. The investigator was concerned that two or more raters might assess students differently. If this were, in fact, the case the differences in rating could effect students' overall scores. Consequently, tests for significant differences (ANOVA) in scores by rater for each station were carried out for the IPS Total scores and for the components of Empathy, Communication, Thoroughness and Satisfaction. In addition, the same tests were carried out on the other competencies of Data Collection, Diagnosis and Management. Where there were significant differences by rater at the .05 level, the individual student's score was modified in the manner described earlier in data analysis on page 30.

Of the 16 stations, IPS Total scores required modification in 10 stations; Empathy scores in 13 stations; Communication scores in 12 stations; Thoroughness scores in 9 stations and Satisfaction scores in 9 stations.

Eleven stations assessed Data Collection skills and of these, scores in 5 stations required modification. None of the 8 stations assessing Diagnosis skills needed their scores modified. Finally, of the 11 stations measuring Management skills, 3 required modification of student scores.

Using the modified scores a mean IPS Total score across all 16 stations was computed for each student. In addition, modified mean Empathy scores, Communication scores, Thoroughness and Satisfaction scores were computed for each student. Finally, modified scores in Data Collection and Management skills were computed for each student. As a result there were two sets of scores: the original scores with rater differences and the modified scores.

The modified scores were then compared with the original set of scores for the whole group of 95 examinees using Pearson's product-moment correlation coefficient and Spearman's rank-order correlation coefficient. It was reasoned that the Pearson correlational analysis would describe the degree to which the two sets of scores varied with each other and the Spearman coefficient would provide

additional information about the degree to which modifying the scores affected students' ranking. The results of these analyses are presented in Table 2 and show exceptionally small variability in the two sets of scores. Appendix F provides scatterplots of the original and modified scores for IPS, Data Collection and Management and shows little variation in the two sets of scores. The coefficient of determination (r^2) indicates that from 90% to 99% of their variation pattern is held in common. In other words, there is very little difference in students' overall scores whether the original or modified scores are used.

TABLE 2

COMPARISONS OF ORIGINAL AND MODIFIED COMPETENCY SCORES USING
CORRELATIONAL COEFFICIENTS (Pearson's r and Spearman's p)

Competency	r^*	r^2	Sp^*
Total IPS	.96	.92	.94
Empathy	.96	.93	.95
Communication	.98	.96	.96
Thoroughness	.99	.98	.99
Satisfaction	.99	.99	.99
Data Collection	.95	.90	.94
Diagnosis #	—	—	—
Management	.99	.98	.98

* p = at least .001 for all correlation coefficients.

Diagnosis competency is not included as no scores required modification.

Research Question 2

The second research question asked "What is the relationship between the interpersonal skill competency and other competencies specifically: (1) Data Collection, (2) Diagnosis, (3) Management and (4) Overall Non-Interpersonal Skill Competency?" To test these relationships, Pearson's product-moment correlations were carried out on the IPS Total and Component scores and the scores for Data Collection, Diagnosis, Management and Overall Non-Interpersonal Skill Competency. Data for these analyses includes scores from all 95 examinees. The results are presented in Table 3. Results for both original and modified scores are given where available.

If the correlations between IPS and Data Collection, Diagnosis and Management are corrected for attenuation because of the inherent unreliability of the data, the correlation coefficients are increased. Specifically, the correlations using original scores become .71 between IPS and Data Collection, .61 between IPS and Diagnosis and .49 between IPS and Management. It was not possible to perform this correction for the relationship between IPS and ON-IPS because the reliability coefficient for the ON-IPS portion of the examination, required in the calculation, was not available.

Overall Non-Interpersonal Competency (ON-IPS) scores included the scores of Data Collection, Diagnosis and Management and, in addition, scores on other competencies such as Test Selection, Test Interpretation, Working Hypotheses, Working Knowledge and Data Interpretation. However, they did not include scores on Communication and Professional Service which had been collected by the University of Manitoba Examination Committee for purposes of comparison of results with Southern Illinois University. The IPS scores correlate moderately to highly across stations with the Communication and Professional Service Checklist scores and it was reasoned that this latter checklist provided a very similar assessment by raters as the IPS. See Appendix D for a table of the correlational coefficients between the IPS scale and the Communication and Professional Service Checklist on 15 of the 16 stations. (The Checklist had not been completed on Station 6). Limitations, related to the way the data had been placed in the data set for analysis, prevented examination of the relationships between ON-IPS and the component skills of IPS.

TABLE 3

RELATIONSHIPS BETWEEN IPS SCORES (Original and Modified¹)
AND OTHER COMPETENCIES, USING CORRELATIONAL COEFFICIENTS

IPS Competency	Data Collection	Diagnosis ²	Management	Overall Non-IPS
Total IPS	.45*** (.57)***	.25* —	.19 (.25)*	.42***
Empathy	.33** (.56)***	.25* —	.09 (.16)	
Communication	.41*** (.53)***	.23* —	.23* (.25)*	
Thoroughness	.44*** (.60)***	.16 —	.25* (.29)**	
Satisfaction	.50*** (.59)***	.25* —	.20 (.21)**	

¹ Correlational coefficients using modified scores appear in brackets.

² Diagnosis scores did not require modification.

* $p < .05$; ** $p < .01$; *** $p < .001$.

Even prior disattenuation, in all cases correlations using the modified scores are somewhat higher than the correlations among the original scores. The relationships are significant at .01 or better for all comparisons involving Data Collection and IPS scores. Most of these relationships are in the low to moderate range. The relationships between IPS scores and Diagnosis and Management are less strong.

After the correlations are corrected for attenuation, the relationships are stronger and reach a level of at least .001 significance.

The correlation between the IPS Total scores and the ON-IPS Competency is .42 and is significant at the .001 level.

To provide more meaningful data for interpretation, Table 4 presents the coefficients of determination for the relationships between these competencies. This shows that the percentages of variation in Diagnosis and Management scores accounted for by the variation in IPS scores is quite small (1% to 8%). However, variation in Total IPS scores can account for 20% to 32% of the variation in Data Collection scores depending on whether original or modified scores are considered. Furthermore, 18 % of the variation in the ON-IPS competency scores is explained by the variation in the IPS scores.

TABLE 4

COEFFICIENTS OF DETERMINATION (r^2) BETWEEN IPS AND OTHER
COMPETENCIES USING ORIGINAL AND MODIFIED SCORES¹

IPS Competency	Data Collection	Diagnosis	Management	Overall Non-IPS
Total IPS	.20 (.32)	.06 —	.04 (.06)	.18
Empathy	.11 (.31)	.06 —	.01 (.03)	
Communication	.16 (.28)	.05 —	.05 (.06)	
Thoroughness	.19 (.36)	.03 —	.06 (.08)	
Satisfaction	.25 (.35)	.06 —	.04 (.04)	

¹ Coefficients using modified scores appear in brackets.

Research Question 3

The third research question asked "Does the interpersonal skill competency vary significantly as a function of the patient problem which the simulated patient presents?" Data from all 95 examinees were included in the analysis for this question. However, it was necessary to exclude an additional station from the analysis as the station involved a simulation of a married couple. The IPS score for each student for this station was the mean of the scores of "the couple". Four simulators (two couples) were trained but during the examination there were cross-overs

between the simulated couples which meant the IPS scores did not involve consistent partners. Consequently, data from only 15 stations are included in the analysis for this question.

A random model of ANOVA was used to analyze the data for this question because the variables: students, raters and stations, constitute only a sample of the total range of possibilities as explained on page 31. Analyses of differences in mean IPS scores between students, between stations and between raters within stations were carried out. The results, presented in Table 5, indicate that there are significant differences in scores in all three instances, that is, between students, between stations and between raters within stations.

TABLE 5

RANDOM EFFECTS ANALYSIS OF VARIANCE SUMMARY TABLE

Source	Sums of Squares	Degrees of Freedom	Mean Squares	F-Ratio
Between Students	30542.30	94	324.92	3.86*
Between Stations	46307.51	14	3307.68	39.29*
Between Raters within Stations	28694.75	20	1434.74	17.04*
Error	109106.94	1296	84.19	
Total	214651.50	1424		

* p = at least .0001.

However, having considered these results, the particular question still remains whether or not significant differences occur between stations which cannot be accounted for by the variability between raters. To answer this question, an F ratio was calculated which compared the variability between stations with the variability between raters within stations. The result of this calculation ($3307.68/1434.74$) reveals that F equals 2.31 with 14 and 20 degrees of freedom.

This F Ratio indicates that there is indeed a difference in IPS scores between stations, significant at

the .05 level, which is greater than can be accounted for by the fact that stations have different raters or simulators.

The mean scores of the stations were then compared to determine which stations differ significantly from each other. A Tukey critical value ($F_t = 4.52$) was used in comparing stations as multiple pairwise comparisons were involved.

Table 6 shows the mean IPS scores of the stations grouped to indicate those that are significantly different from others along with a brief descriptor of each station. Stations within brackets are not significantly different from each other. There are no immediately obvious or common features to stations within groupings which might explain their relative IPS scores.

TABLE 6
MEAN IPS SCORES BY STATION, GROUPED BY SIMILARITY

Station Name	Description	Mean IPS
Larson (#8)	26 yr woman, pelvic pain	56.3
Taylor (#4)	22 yr woman, paralyzed legs	57.4
Juranek (#10)	65 yr woman, jaundice 3rd hospitalization	57.6
McCray (#14)	42 yr man, abdominal pain	61.0
Furlow (#18)	33 yr man, high blood pressure	61.2
Simcoe (#13)	19 yr woman, panic attacks	63.1
Lawrence (#3)	39 yr man, low back pain	65.2
Hansen (#15)	5 yr boy, short stature with 29 yr mother	66.1
Jones (#19)	70 yr woman, memory loss, with husband	66.2
Strawson (#12)	67 yr man, dizziness	68.2
Caressell (#6)	25 yr mother of baby who fell	68.0
Marchetti (#2)	75 yr woman, pre-op assessment	68.9
Towns (#1)	68 yr man short of breath	71.7
Kenny (#16)	26 yr woman, rt. side weakness	72.9
Hutton (#20)	65 yr woman, fatigue	75.9
Mean Across Stations		65.4

Spearman's rank-order correlations were computed for the age of the patient and the IPS score. These results show $Sp = .43$ or $.41$ depending if the the age of the mother or child respectively is used in station 15. This comparison does not reach significance ($t = .48$; $df = 13$). Comparisons were also made between the mean station IPS scores and the overall station competence scores ($Sp = .11$) and between the simulated patient's age and overall competence scores ($Sp = .33$ or $.20$, depending on the age used in station 15). None of these are significant relationships.

Research Question 4

The fourth research question asked "What is the predictive accuracy of the medical students at assessing their own interpersonal skills?" A predictive accuracy score was obtained for each student as described in the Method chapter. A smaller score indicates a greater degree of accuracy on the part of the examinees at predicting how the simulator had assessed them.

Of the 95 individuals who took the examination, 62 (65%) predicted a lower score for themselves (Under Raters) than the simulator gave them. The mean IPS score for this group was 65.56 (S.D. 4.62). The remaining 33 examinees (35%), the Over Raters, predicted the simulators would rate them better than actually occurred. The mean IPS score for

this group was 65.12 (S.D. 4.53). The mean scores of the groups were compared using ANOVA and found not to be significantly different ($F = .20$; $df 1,93$; $p = .57$).

A further analysis was conducted on the IPS scores of those individuals whose predictions were more than 20 points different from what the simulators had assessed. Eighteen Under Raters (UR's) came into this category and 5 Over Raters (OR's). The mean IPS scores of these sub-groups are: UR's = 65.40 (S.D. 3.89) and OR's = 61.39 (S.D. 7.24). These means were compared using ANOVA and found not to be significantly different ($F = 2.84$; $df 1,21$; $p = .10$).

The predictive accuracy score for each individual was compared to the mean IPS score for that individual using Pearson's Product Moment Correlation. These scores were found to be negatively correlated at $-.21$. This relationship is significant at the $.05$ level and suggests that higher IPS scores tend to be associated with greater ability to predict one's interpersonal skills. The predictive accuracy scores were further compared with scores on the IPS components of Empathy ($r = -.10$; ns) and Communication ($r = -.32$; $p < .01$). Of the two comparisons only the relationship with Communication scores is at the significant level.

Supplementary Findings

Although not one of the original research questions, a factor analysis was carried out on the 14 items of the IPS scale to determine whether or not the scale appeared to measure the four components of Empathy, Communication, Thoroughness and Satisfaction. The data for factor analysis came from 15 of the 16 stations. One station was excluded because the variance of the data was near zero, indicating that there was not much variation in raters' assessments of individual scale items. Using the usual criterion of an eigenvalue of 1, in fact, only two factors were found accounting for 72.62% of the variation. Table 7 presents the eigenvalue summary for the IPS scale.

The next step was to determine how each of the scale items was loaded on the two identified factors and then to employ a varimax rotation to provide for more meaningful interpretation of these loadings. The rotated factor loadings are presented in Table 8. The IPS scale items are identified by the component they were purported to measure when the scale was developed. The complete items can be referred to in the IPS Rating Scale in Appendix A.

TABLE 7

EIGENVALUE SUMMARY IPS RATING SCALE

Factor	Eigenvalue	Percent	Cumulative Percent
1	8.7379	62.41	62.41
2	1.4285	10.20	72.62
3	0.7949	5.68	78.29
4	0.6883	4.92	83.21
5	0.5221	3.73	86.94
6	0.4503	3.22	90.16
7	0.3402	2.43	92.59
8	0.2093	1.50	94.08
9	0.1780	1.27	95.35
10	0.1659	1.19	96.54
11	0.1423	1.02	97.55
12	0.1220	0.87	98.43
13	0.1130	0.81	99.23
14	0.1074	0.77	100.00

TABLE 8

ROTATED FACTOR LOADINGS IPS RATING SCALE

IPS Item	Factor 1	Factor 2	Communality
Empathy 1	0.4914	0.7322	0.7776
Empathy 2	0.0966	0.7985	0.6470
Empathy 3	0.5151	0.7162	0.7783
Empathy 4	0.4485	0.7791	0.8082
Empathy 5	0.4878	0.7120	0.7449
Empathy 6	0.3250	0.8206	0.7791
Communi- cation 1	0.7035	0.4737	0.7193
Communi- cation 2	0.7847	0.3655	0.7494
Communi- cation 3	0.8690	0.1628	0.7817
Communi- cation 4	0.8582	0.3425	0.8539
Communi- cation 5	0.1314	0.6688	0.4646
Communi- cation 6	0.7654	0.1092	0.5978
Thoroughness	0.6693	0.4051	0.6121
Satisfaction	0.7890	0.4794	0.8525

Interpreting factor loadings is a matter requiring considerable judgment as well as the use of mathematical propositions. The factor loadings can be considered as correlational coefficients and the higher the coefficient the more strongly the item loads on a specific factor. Only coefficients $> .30$ are considered to be significant loadings, although this rule of thumb has more relevance

when the number of items being analyzed is more than 50. Nevertheless, in examining factor loadings in Table 8, Factor 1 appears to correspond most closely to the Communication items on the IPS rating scale. Communication item #5 is the exception. The Thoroughness and Satisfaction items load most strongly on Factor 1 as well and do not load as separate factors. As well, several of the Empathy items load moderately on Factor 1. However, all the Empathy items load most strongly on Factor 2 which includes, in addition, Communication item #5 : the language the student doctor used. However, over 50% of the variance on Communication item #5 remains unaccounted for as indicated by the Communality column on the right. The Communality column indicates the variance of the item accounted for by the two factors and can be viewed as a percentage.

The 5 items of the Communication and Professional Service Checklist from Southern Illinois University were then included in the factor analysis along with the 14 IPS items. This was done to determine if the Checklist items measured additional factors to that measured by the IPS scale. When an eigenvalue of 1 is used, the 19 items load on 3 factors accounting for 77.24% of the variation. These results are presented in Table 9.

TABLE 9

**EIGENVALUE SUMMARY IPS RATING SCALE AND
COMMUNICATION AND PROFESSIONAL SERVICE CHECKLIST**

Factor	Eigenvalue	Percent	Cumulative Percent
1	11.9046	62.66	62.66
2	1.6553	8.71	71.37
3	1.1158	5.87	77.24
4	0.7653	4.03	81.27
5	0.6487	3.41	84.68
6	0.4843	2.55	87.23
7	0.4279	2.25	89.48
8	0.3450	1.82	91.30
9	0.2772	1.46	92.76
10	0.2467	1.30	94.06
11	0.1869	0.98	95.04
12	0.1726	0.91	95.95
13	0.1584	0.83	96.78
14	0.1363	0.72	97.50
15	0.1236	0.65	98.15
16	0.1115	0.59	98.74
17	0.1035	0.54	99.28
18	0.0896	0.47	99.76
19	0.0465	0.24	100.00

The loadings of the 19 items on the three factors were computed and then rotated by varimax rotation to provide more meaningful interpretation. The rotated factor loadings are presented in Table 10. For a clearer understanding of the content of each of the 19 items, the IPS Rating Scale and the Communication and Professional Service Checklist can be referred to in Appendices A and B respectively.

Applying the same guidelines and reasoning to this factor analysis as was done in the case of the IPS items alone, here items which pertain to receiving an explanation about one's medical condition and treatment appear to load most strongly on Factor 1. In fact, this factor could generally be termed Getting an Explanation. Only Empathy items load significantly on Factor 2 although two of those (#3 and #5) also load strongly on Factor 3. Nonetheless, the second factor could be termed Empathy. The third factor has predominantly elements which might be described as Professional Behaviour, and includes the doctor's personal approach and professionalism and the type of questions and language used. Language also loads moderately on Factor 2 but almost 50% of the variance of Communication #5, (language used) remains unaccounted for as shown by the Communality. Thoroughness and Satisfaction again do not load as separate factors but load fairly evenly on both Factors 1 and 3. However, there is still 40% of the variance related to Thoroughness not accounted for by the three factors.

TABLE 10

ROTATED FACTOR LOADINGS IPS RATING SCALE AND
COMMUNICATION AND PROFESSIONAL SERVICE CHECKLIST

Item	Factor 1	Factor 2	Factor 3	Communality
Comm 1	0.3170	0.1432	.8444	0.8341
Comm 2	0.7715	0.1443	.4501	0.8185
ProfSer 1	0.4266	0.1109	.7530	0.7613
ProfSer 2	0.4006	0.3381	.6413	0.6861
ProfSer 3	0.6168	0.2394	.6779	0.8973
Empathy 1	0.4088	0.6906	.3929	0.7984
Empathy 2	0.1518	0.8768	.0277	0.7925
Empathy 3	0.3356	0.5805	.5882	0.7956
Empathy 4	0.3967	0.7691	.3070	0.8431
Empathy 5	0.2632	0.5171	.6920	0.8156
Empathy 6	0.2544	0.8006	.3178	0.8067
Communi- cation 1	0.5504	0.3530	.5501	0.7301
Communi- cation 2	0.7112	0.3189	.3720	0.7439
Communi- cation 3	0.8170	0.1589	.2837	0.7732
Communi- cation 4	0.7835	0.3261	.3519	0.8440
Communi- cation 5	-.0470	0.4679	.5553	0.5295
Communi- cation 6	0.8145	0.2098	.0494	0.7099
Thoroughness	0.5109	0.3055	.4885	0.5930
Satisfaction	0.6091	0.3119	.6578	0.9010

Discussion

Research Question 1

If overall scores are considered most of the University of Manitoba graduating students were assessed in terms of their interpersonal skills in a positive way by the simulated patients. If a score two standard deviations below the mean is used as a minimally acceptable standard, then only two students scored below this and another four students scored close to this. While this seems to suggest that University of Manitoba graduating students are competent in this area, there is no actual gold standard by which to judge them.

What is an acceptable level of performance in interpersonal skills for graduating medical students? Is it appropriate to judge interpersonal skills in this norm-referenced manner? In some way these few students "stood out" from their peers in interpersonal skills. Until there is a gold standard, such a measure of difference has some relevance in considering acceptable performance. How reasonable are the expectations of simulated patients? It is this investigator's impression from many conversations with simulated patients, that if simulated patients have doubts in making decisions about students' performance, they tend to err on the side of the student and to give them the "benefit of the doubt."

Looking at the results somewhat differently and considering the maximum score that a simulator could have given a student as 84, then these 6 students demonstrated less than a 70 % level of interpersonal skill.

It does seem reasonable to conclude that the six students noted above were considered by the simulated patients to be lacking to some degree in interpersonal skills. What is important to emphasize is that through this assessment, individuals with deficiencies in the interpersonal area can be identified. Remediation can then be planned and implemented prior to their graduation.

The students were generally considered empathetic but the skill level of those who were rated lower than most of their classmates is of concern from the patients' point of view. Three students are assessed as below two standard deviations from the mean. When their scores are considered in relation to a possible maximum, they are viewed as demonstrating less than 68% of the skills of Empathy. It appears that these students are perceived as not understanding the patient's point of view or perhaps even not making an effort to do so. Allowing the patient to tell his or her own story is linked to satisfaction and compliance (Stewart, 1984; Stiles et al., 1979) and thereby, would seem a crucial skill for all physicians. Skills of demonstrating empathy are perhaps more difficult to teach

than specific techniques of communication. Remediation, however, is possible (Grayson, Nugent & Oken, 1977).

Similar comments can be made about the skills of the three students with the lowest scores in the Communication component. These individuals are viewed as making less effort than their peers to provide explanations and information about the patient's illness and treatment. Important as such skills are for compliance, the results suggest that remediation should take place for certain students. Remediation might initially involve demonstrating to them the value of information sharing and then teaching them specific techniques of sharing information with patients.

The scores on the single items of Thoroughness and Satisfaction suggest that most students, as assessed by simulators, demonstrated appropriate levels of skill. Two students were below two standard deviations from the mean in Thoroughness and demonstrated less than 52% of the maximum Thoroughness score. In Satisfaction, four students fell below this level and demonstrated less than 66% of the maximum Satisfaction score. While specific skills of Thoroughness and Satisfaction are not, in themselves teachable, efforts directed to improving skills of empathy and communication might indirectly improve the perception of student thoroughness and increase the simulated patients' sense of satisfaction with the encounter.

In summary, the results suggest that most of the University of Manitoba students displayed an acceptable level of interpersonal skills. For the few students who appear to be somewhat lacking in this area, the issue of remediation is raised. The Clinical Comprehensive Evaluation for graduating students occurs approximately five months prior to students' anticipated graduation. This time should be adequate to modify student behaviour in interaction with patients unless there is an underlying attitude problem.

The variability found in scores of raters within stations is of concern to someone who organizes a performance-based examination. Such variability may have consequences for student scores if IPS competency is evaluated on a station by station basis. If stations utilize more than one simulator as rater, then variability between raters should be analyzed when examination results are being compiled. If significant differences are found between raters' scores, student scores should be modified as was done in this study. Extensive training of simulators in assessing students could serve to minimize such differences in rating.

However, if IPS competency is considered across the total number of stations, then the findings of this investigation have other relevance. The high and positive correlation between the raw scores and the modified scores indicates that while there may be differences between raters

on individual stations, if an adequate number of stations is used, (about 15 to 16), such differences become inconsequential for overall student scores. It then becomes less critical that extensive time is spent in training simulated patients for inter-rater reliability. In the real world where time and monies for training are usually limited, and when practical choices must be made, time spent on training simulators to be accurate and consistent in their role performance is likely a better use of limited resources. This is not to suggest that simulators do not need some training in assessment of students. Training should at the minimum consist of: (1) a specific explanation of the rating scale items and (2) at least two opportunities to practice assessing individuals using the rating scale followed by a debriefing. If more than one simulator is being used per station, then the training and practice should occur jointly to allow for sharing of questions, problems and viewpoints. However, attaining .80 or greater inter-rater reliability is not crucial, if students are to be assessed an adequate number of times.

Research Question 2

Considering the correlation coefficients prior to disattenuation, the moderately high and positive correlation found between IPS scores and Data Collection scores and the weaker relationship found between IPS scores and Diagnosis scores (Table 3) are consistent with those found by Stillman

and her colleagues (1986). Intuitively, these results can be interpreted to indicate that effective interpersonal skills enable physicians to collect appropriate information from patients. Effective interpersonal skills encourage revelation of the patient's story and thereby help the physician to develop a better understanding of the problem. Despite the statistical significance of these relationships, the coefficients of determination (Table 4) indicate that there is a considerable amount of the variation in the Data Collection scores and even more of the Diagnosis scores not determined by the effectiveness of interpersonal skills.

The strong relationships found between Data Collection and the components of Thoroughness and Satisfaction may be exaggerated due to the special knowledge of a simulator and may not be reflective of patients' real life responses. Through training and experience with grading students' performance, simulators become aware of what is expected of students. Although they are specifically instructed not to assess students interpersonally in relation to the number of items obtained on the Data Collection checklists, some simulators on some occasions may do this. They may have based their perception of the student's thoroughness and their own satisfaction on the amount of information obtained by the student.

The stronger relationships found between IPS scores and Data Collection and Diagnosis scores found after the

correlations are corrected for attenuation, lends support to the previously stated interpretation that effective interpersonal skills assist the physician to collect information from the patient and then to arrive at a diagnosis.

The weaker relationship between IPS scores and Management scores shown in Table 3 is not surprising. The assessment of Management skills in this examination was generally an assessment of students' knowledge of appropriate management rather than of their ability to implement a management plan in negotiation with a patient. The results seem consistent with the lack of relationship found between knowledge and interviewing skills by researchers Kraan and Crijen (cited in Stewart, Brown & Weston, 1988) and by Wolraich, Albanese, Reiter-Thayer and Barratt, (1982). The results of the Certification Examinations of the College of Family Physicians of Canada have similarly shown only weak relationships between management scores and affective skill scores (P. Rainsberry, personal communication, April 10, 1989).

However, the relationship between Management scores and IPS scores becomes stronger after disattenuation. Other researchers have not reported correlations corrected for attenuation and therefore, it is not possible to compare the results.

The significant, although not strong, relationship found between IPS scores and the ON-IPS scores is somewhat different than the findings of Henbest (cited in Stewart et al., 1988) who claims that there is virtually no correlation between patient-centredness and level of medical competence. While the present study did not measure patient-centredness specifically, the items on the IPS scale were very similar to items assessing patient-centred attributes. It would seem in relation to the University of Manitoba examinees, at least, there is a relationship between being competent interpersonally and being competent in other medical skills.

It is not unreasonable to assume that when students know how to deal with a medical problem, a sense of confidence carries over into the interaction with the patient and affects to some extent their interpersonal effectiveness.

Research Question 3

The finding of significant differences in IPS scores of stations not accounted for by the variability in raters is consistent with other research which has shown that assessed competence is case specific (Norman, Neufeld, Walsh, Woodward & McConvey, 1985; Stillman et al., 1986; Williams et al., 1987). This confirms the need to obtain a broad sample of student behaviour by assessing students over an adequate number of stations which in the case of interpersonal skills

is approximately 15. With 15 cases, the reliability coefficients in this area of interpersonal skills are about .80 which is considered acceptable (Stillman et al., 1986).

Despite the fact that the relationship between age and IPS score did not reach significance, this area seems worthy of further investigation. Possibly, medical students feel less comfortable asking personal questions of patients in their peer group and this discomfort is translated into less effective interpersonal skills. Typically, the clinical experience obtained by most medical students is with older patients who are either admitted to hospital or attended to in clinics. Possibly it is their relative inexperience at dealing with peers as patients that results in such feelings of discomfort.

The three stations with the lowest IPS scores do not deal with the same medical content but may all have presented an interpersonal challenge for the medical students. Station 8 required the students to inquire for personal information regarding the patient's sexual and marital relationship. This investigator's teaching experience suggests that medical students are often uncomfortable when asking such personal questions. This observation has some empirical support. Maguire & Rutter (1976) in a study of senior medical students found that more than one-third of them avoided asking any specific questions about their patients' personal relationships, including

their sexual adjustment or feelings about their illnesses. Station 4 required the students to interview a distressed young woman, with paralyzed legs. This situation, too, might have made students feel uncomfortable. Finally, Station 10 presented the students with an older woman who had been hospitalized three times and now was frustrated at not getting better. Dealing with these negative emotions and the challenge of having to solve a problem unsolved in two previous hospital admissions may have made the students uncomfortable. Discomfort or a lack of confidence may lead to less effective interpersonal skills.

If the situations described above, actually made the students feel inordinately uncomfortable, and if this discomfort led to their being less effective interpersonally, then these findings have special implications for medical education. In their future practices medical graduates will encounter, on a daily basis, patients of whom they must ask deeply personal questions and these patients will express feelings of anger, frustration, fear and anxiety. It is important that physicians develop comfort in dealing with such clinical situations. Learning effective interpersonal skills requires that students be given an opportunity to practice them under observation and to receive feedback. This learning will be most relevant if carried out during the clinical years as well as during a pre-clinical course. At present, there is

not a lot of opportunity for students to have their history-taking and physical examination skills directly observed by faculty during their clinical experience.

Research Question 4

Conclusions from the findings in regard to this question, must be presented with caution as they pertain to only one encounter and one simulator. However, many of these examinees appear to have difficulty judging how a simulator felt about their interpersonal skills. While the predictive accuracy score is not a direct measure of their self-assessment skills, it is likely a good indication of how they perceive their interpersonal skills. Almost one-fifth of the examinees underrated their skills by 20 or more points which might suggest a certain lack of confidence on their part and a definite inability to judge how they are coming across to others. Do they always under rate their skills in this area or might their inability to judge lead them to be unaware at other times when there are problems in the interaction with a patient? The answer is not clear from the results. Perhaps of more concern are those 5 individuals who overrated how the simulator was assessing them. At least in this one encounter, they did not detect that the simulated patient was dissatisfied with the interaction. Does this lack of awareness occur in other patient situations? More emphasis placed on developing self-

awareness during training in interpersonal skills might benefit these students.

The relationship between predictive accuracy and IPS level shows that those individuals with better interpersonal skills are probably more aware of how others feel about them. These students likely have a certain amount of intuition which helps them to read others reactions and enables them to be more sensitive to others' worries and need for explanation. This finding is consistent in a general way with what Harman (1986) found. However, the subsequent finding of a non-significant relationship between the predictive accuracy score and Empathy specifically does not substantiate Harman's findings. Instead the present study shows a relationship between predictive accuracy and the Communication skills component. Harman did not measure communication and measured empathy more extensively and these differences may explain the difference in specific findings.

As discussed in Chapter 1, the ability to be aware when one is encountering problems in an interaction with a patient, is important for the practicing physician in order to correct miscommunication and misunderstanding. Further investigation of students' self-assessment abilities in the area of interpersonal skills seems warranted.

Supplementary Findings

The factor analysis of the IPS Rating Scale suggests that there were two main components measured. Factor 1, accounting for 62% of the variation in scores, seems to consist of items measuring the degree to which information is shared with the patient and how much the patient is involved in the management process (Tables 7 and 8). This finding indicates that in this investigation the most significant component of interpersonal skills was that related to communication of information and patient participation. Neither Thoroughness nor Satisfaction were analyzed as separate factors. Both load significantly on the two factors suggesting patients' attitudes to these items are determined by other features of the encounter. The most significant features are those related to sharing of information (Factor 1). However, Thoroughness and Satisfaction do load to a lesser degree on Factor 2 with measures of Empathy. This loading suggests that a feeling of being understood also features in patients' impression of thoroughness and contributes to a sense of satisfaction.

All the Empathy items load most strongly together on Factor 2 but account for only 10% of the variation. In many cases there are also moderate loadings on Factor 1 suggesting that feeling understood and effective verbal communication are hard to separate. Understandable language (Communication #5) also loads most strongly with Empathy

rather than Communication which fact does not seem immediately explainable. It is possible that patients regard the use of understandable language as an effort by the physician to understand them. Furthermore, in that only 46% of the variation of this item is accounted for by the two factors, type of language used may be a distinctive item.

The factor analysis of the combined IPS Rating Scale and the Communication and Professional Service Checklist (Tables 9 and 10) suggests that there may be additional aspects related to patient attitudes that are assessed on the latter checklist which are omitted when the IPS Rating Scale is used alone. These missing aspects involve the patient's perception that the physician's approach is organized and conveys a competent and courteous manner. Items which would measure these aspects of the doctor's professional approach should be included in amendments to the IPS Rating Scale.

CHAPTER V

CONCLUSIONS, LIMITATIONS AND RECOMMENDATIONS

The purpose of this investigation was to assess the interpersonal skills of the 1988 class of graduating medical students at the University of Manitoba. This assessment was conducted as part of a 20 station performance-based examination, administered to the students in December, 1987. Nineteen of the stations were "live" simulations in which standardized patients were trained to present the signs and symptoms of typical medical problems. The standardized or simulated patients assessed the interpersonal skills of the students using a rating scale designed for that purpose. Because such an assessment of graduating medical students had not been completed previously and because there were no established norms available for comparison, this investigation is considered exploratory in nature.

Specifically, the investigation asked four research questions:

1. What is the interpersonal skill competency level of the graduating medical students, as a group, at the University of Manitoba?

2. What is the relationship, if any, between the level of interpersonal skill competency and other competencies,

specifically: (1) data collection (2) diagnosis (3) management and (4) overall non-interpersonal skill competency?

3. Does the level of interpersonal skill competency vary significantly as a function of the patient problem which the simulated patient presents?

4. What is the predictive accuracy of the medical students at assessing their own interpersonal skill competency?

The review of the literature indicated a number of important points:

1. Effective interpersonal skills are associated with increased patient compliance and satisfaction. While interpersonal skills are taught to medical students early in their training, there is doubt that the skills are retained by the time of graduation.

2. Two important components of interpersonal skills are those related to: (1) demonstrating empathic understanding towards the patient and (2) providing effective explanations to patients about their illness and plans for follow-up.

3. Employing standardized patients to assess the interpersonal skills of graduating medical students is practical, feasible and appropriate.

4. Moderately high positive correlations exist between interpersonal skill competency and competency in skills of collecting information from patients.

5. Some evidence indicates that interpersonal skills, like many other medical skills, are context specific and are related to the patient problem.

6. Findings from the field of psychology suggest that individuals with higher levels of empathy are more likely to predict accurately how others perceive them, than individuals with lower levels of empathy.

A mean score was computed for each student's interpersonal skill performance at each station and for all the stations combined. All stations had at least two raters. When a significant inter-rater difference occurred in assessing students, as determined by ANOVA, student scores were modified to reflect that difference. This resulted in two sets of scores: the original scores and the modified scores. These scores were then compared using correlational analyses to determine the extent to which the difference in raters affected the overall student scores. IPS competency was compared to other medical competencies using Pearson's product-moment correlations. A random model of ANOVA was used to determine whether or not there was a difference between stations which was greater than that accounted for by the variability in raters. Finally, a predictive accuracy

score was computed for each student and these scores were compared to the interpersonal skill competency scores using Pearson's product-moment correlations. The Number Cruncher Statistical System was utilized with a personal computer to analyze the data.

Conclusions

This study revealed that, in general, the 1988 class of graduating medical students at the University of Manitoba, were assessed favourably by the simulated patients in regard to their interpersonal skills. A few students, however, were identified as somewhat deficient interpersonally in interactions with the standardized patients. The investigation confirmed that the students' overall IPS scores were not affected significantly by the use of different raters. The IPS scores showed a moderate and positive correlation with skills of gathering information from patients and with overall competency in non-interpersonal skill areas. Some evidence indicated that IPS scores varied significantly across the stations, to a greater extent than was accounted for by rater variability, suggesting that the nature of the patient problem has an effect on the student's interpersonal skill competence. Students with higher IPS scores were more likely to predict accurately how a simulator would assess their interpersonal skills than students with lower scores.

Considered generally these results provide reassurance to those who teach interpersonal skills to medical students. However, the results are somewhat more favourable than might be anticipated from a review of the literature, which suggests that many graduating medical students are deficient in interpersonal skills. If the results of this present study are reliable and valid, it seems reasonable to conclude that the program for teaching interpersonal skills at the University of Manitoba is particularly effective. However, the stations which formed the examination varied in their interpersonal challenge to the students and it may be that a different set of cases would have shown quite different results. The study has shown that it is important to consider both the content of the medical problem being presented and its potential for interpersonal challenge when selecting examination cases. Cases do not provide equal interpersonal challenge and are, therefore, not interchangeable for assessing interpersonal skills. This study offered some evidence that the cases on which the students performed least well, were those which may have presented a particular interpersonal challenge to the students. This investigation has provided moderately strong evidence that competence in interpersonal skills is associated with competence in certain other medical skills, most specifically, skill in gathering information from patients. This finding confirms the view that effective interpersonal skills are highly relevant in enabling the

physician to develop a more complete understanding of the patient's problem. The study also suggests that medical students with more effective interpersonal skills may be able to avoid breakdowns in communication with patients because they are able to perceive when a patient is regarding their interaction unfavourably.

Limitations

In light of certain limitations, the results of this study must be viewed with caution. The simulated patients were required, due to the examination schedule, to assess the students' interpersonal skills in usually less than 5 minutes between interactions. They often admitted to feeling pressed for time and their ability to discriminate between the effectiveness of students may have been affected and may have led to some inaccuracy in the assessment of interpersonal skills. In addition, some simulators stated that they found the format of the rating scale, specifically the range of scores from +3 to -3, to be awkward. Therefore, the format might also have affected simulator accuracy. Finally, the simulated patients did not receive the extent of training in assessing students that the investigator had originally planned. An increase in training, some modification to the rating scale and more time between interviews might all contribute to greater accuracy in grading. At one station, the lack of variance in the data which became evident during factor analysis may be the

result of both inadequate training and limited time for assessment.

Further consideration of this possible limitation suggests that if individual simulators did have problems with assessing some students accurately, this situation does not appear to have had strong effects on the students' overall IPS scores. The fact that there was very little difference between the original and modified overall student scores provides evidence for this conclusion.

The limitation on time between interactions also meant that a form, more abbreviated than that originally planned for, was used. As a result the attempt to measure empathy displayed by the students was less than optimal and consequently, the findings in regard to empathy must be considered cautiously.

The study assessed students' IPS in an examination setting where they knew the simulators were grading their performance. One might argue that these students would not perform at the same level with actual patients in a clinical setting. It seems that such an argument has not been validated by other empirical evidence. Research has shown that performance with simulated patients is not significantly different than performance with actual patients (Norman et al., 1982; Sanson-Fisher & Poole, 1980).

An additional limitation might be posed that this study did not assess students' interpersonal skills over an extended period of time with any one patient. It might be argued that it is perhaps over a longer time period, when the physician and patient must negotiate treatment plans and deal with possible differences of opinion, that deficits in interpersonal skills are more likely to become evident. It is in such interactions that effective interpersonal skills become most critical. Unquestionably, the interpersonal skills required for such long term relationships are more difficult to assess in a performance-based examination format as described here for a single encounter. However, a suggestion for how such an assessment might be accomplished is described in Research Recommendations.

Recommendations

The data collected here seems to provide strong evidence that the program within the Faculty of Medicine at the University of Manitoba for teaching interpersonal skills should continue. As a result of this study the following specific recommendations can be made:

1. The literature suggests and this study further confirms that interpersonal skill competency is related to the content of the medical problem. Therefore, it is important that students have extended experience with a wide variety of patient problems. Furthermore, it is highly

desirable that students receive relevant and immediate feedback about their performance in these encounters.

2. Particular efforts should be made to offer experience and feedback to students in dealing with "difficult" patient encounters which provide a challenge concerning interpersonal communications. Such learning opportunities, given more time and resources, are possible.

3. Further efforts should be made to identify, early in their training, students who demonstrate deficits in interpersonal relationships with patients. The increased observation and feedback required to implement the previous two recommendations would likely result in such early identification. Remedial programs could then be undertaken to assist these students to improve their interpersonal skills with patients.

4. Self-assessment of interpersonal skills should be included in the teaching program. Many students were quite inaccurate in predicting how the simulators would assess them. Supervised guidance in self-assessment of interpersonal skills would go far toward helping medical students to avoid miscommunications in future relationships with patients.

5. The assessment of interpersonal skills of graduating medical students at the University of Manitoba should continue. Assessment of interpersonal skills should be part

of any performance-based examination which is held and the results should be compared from one year to the next to establish local norms.

Research Recommendations

1. More research is needed to identify the interpersonal skill competence demonstrated by medical students and more senior practitioners across a broad range of patient problems. Such investigations are required to establish which medical problems presented by standardized patients provide comparable degrees of interpersonal challenge. Once these results are obtained, the process of planning the content of future examinations to assess interpersonal skills would be more objective and consequently, provide a more effective assessment of interpersonal skills.

2. Consideration should be given to designing stations that provide particular interpersonal challenges for use in the Clinical Comprehensive Examination for graduating medical students. Examples of such stations might be the "non-compliant" patient, the "angry" patient, the patient who is terminally ill, the patient with multiple problems or the patient who cries. The Certification Examination of the College of Family Physicians of Canada, which also employs simulations, consists of one or more stations requiring the candidate to deal with a "difficult" patient in a single 15 minute interview. All of the "difficult" encounters listed

above will likely be faced by the students after they enter practice.

3. Further research is needed to measure the medical students' accuracy in self-assessment of interpersonal skills over several encounters with simulated patients. It is very important to on-going therapeutic relationships that physicians be able to perceive when these relationships are "getting into trouble".

4. In view of the within station variability in raters' assessments, consideration should be given to using one simulator per station where possible. However, research is needed to identify the number of times a simulator is able to perform a simulation and still maintain accuracy and consistency in both role performance and assessment of students.

References

- Abrahamson, S. (1985). Assessment of student clinical performance. Evaluation and the Health Professions, 8, 413-427.
- Barrett-Lennard, G.T. (1962). Dimensions of therapist response as causal factors in therapeutic change. Psychological Monographs, 76, (Whole No. 562).
- Barrett-Lennard, G.T. (1978). The relationship inventory: Later development and adaptations. JSAS Catalog of Selected Documents in Psychology, 8, 68.
- Barrett-Lennard, G.T. (1981). The empathy cycle: Refinement of a nuclear concept. Journal of Counseling Psychology, 28, 91-100.
- Barrows, H.S. (1971). Simulated patients: The development and use of a new technique in medical education. Springfield, IL: Charles C. Thomas.
- Byrne P.S. & Long, B.E. (1976). Doctors talking to patients. London: Her Majesty's Stationary Office.
- Carkhuff, R.R. & Berenson, B.G. (1967). Beyond counseling and therapy. New York: Holt, Rinehart & Winston.

- Carroll, J.G. & Monroe, J. (1980). Teaching clinical interviewing in the health professions: A review of empirical research. Evaluation and the Health Professions, 3, 21-45.
- Child, D. (1970). The essentials of factor analysis. London: Holt, Rinehart & Winston.
- Claiborn, C.D., Crawford, J.B. & Hackman, H.W. (1983). Effects of intervention discrepancy in counseling for negative emotions. Journal of Counseling Psychology, 30, 164-171.
- Coulehan, J.L. & Nardini, D. (1982). Patient simulation and team-teaching in a medical interviewing course. Journal of Medical Education, 57, 560-562.
- Cousins, N. (1985). How patients appraise physicians. The New England Journal of Medicine, 313, 1422-1424.
- 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.

- English, H.B. & English, A.C. (1958). A comprehensive dictionary of psychological and psychoanalytical terms. New York: Longmans, Green.
- Faber, R.J., Out, J. & Reepmaker, J. (1984). Teaching interviewing skills to paediatric junior clerks using simulated patients. Medical Education, 18, 255-261.
- Felletti, G.I. & Carney, S.L. (1984). Evaluating patients' satisfaction with medical students' interviewing skills. Medical Education, 18, 15-20.
- Ferguson, G.A. (1971). Statistical analysis in psychology and education. New York: McGraw Hill.
- Fisher, A.W. (1971). Patients' evaluation of outpatient medical care. Journal of Medical Education, 46, 238-244.
- Francis, V., Korsch, B.M. & Morris, M. J. (1969). Gaps in doctor-patient communication: Patients' response to medical advice. The New England Journal of Medicine, 280, 535-540.
- Froelich, R.E. (1969). A course in medical interviewing. Journal of Medical Education, 44, 1165-69.

- Grayson, M., Nugent, C. & Oken, S.L. (1977). A systematic and comprehensive approach to teaching and evaluating interpersonal skills. Journal of Medical Education, 52, 906-913.
- Hannay, D.R. (1980). Teaching interviewing with simulated patients. Medical Education, 14, 246-248.
- Harden, R.M. & Gleeson, F.A. (1979). Assessment of clinical competence using an objective structured clinical examination (OSCE). Medical Education, 13, 41-54.
- Harman, J.I. (1986). Relations among components of the empathic process. Journal of Counseling Psychology, 33, 371-376.
- Harper, A.C., Roy, W.B., Norman, G.R., Rand, C.A. & Feightner, J.W. (1983). Difficulties in clinical skills evaluation. Medical Education, 17, 24-27.
- Hassard, T.H., Kopelow, M., Schnabl, G. & Klass, D.J. (1988). Performance and perception in a simulated clinical encounter. Proceedings of the Twenty-Seventh Annual Conference, Research in Medical Education, 38-43.
- Hintze, J.L. (1986). Number Cruncher Statistical System (Statistical Computer Package). 865 East, 400 North, Kaysville, Utah, 84037.

- Holmes, F.F., Baker, L.H., Torian, E.C., Richardson, N. K., Glick, S. & Yarmat, A.J. (1978). Measuring clinical competence of medical students. Medical Education, 12, 364-368.
- Irwin, W.G. & Bamber, J.H. (1984). An evaluation of medical student behaviours in communication. Medical Education, 18, 90-95.
- 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.
- 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.
- Kauss, D. R., Robbins, A.S., Abrass, I., Bakaitis, R.F. & Anderson, L.A. (1980). The long-term effectiveness of interpersonal skills training in medical schools. Journal of Medical Education, 55, 595-601.
- Kim, J.O. & Mueller, C.W. (1978). Factor analysis: Statistical methods and practical issues. Beverly Hills, CA: Sage.
- Knox, J.D.E., Alexander, D.W., Morrison, A.T. & Bennett, A. (1979). Communication skills and undergraduate medical education. Medical Education, 13, 345-348.

- Korsch, B.M., Gozzi, E.K. & Francis, V. (1968). Gaps in doctor-patient communication: I Doctor-patient interaction and patient satisfaction. Pediatrics, 42, 855-871.
- Mace, D.R. (1971). Communication, interviewing, & the physician-patient relationship. In R.H. Coombs & C.E. Vincent (Eds.), Psychosocial aspects of medical training. Springfield, IL: Charles C. Thomas.
- Maguire, G.P., Clarke, D. & Jolley, B. (1977). An experimental comparison of three courses in history-taking skills for medical students. Medical Education, 11, 175-182.
- Maguire, P., Fairbairn, S. & Fletcher, C. (1986). Consultation skills of young doctors: II Most young doctors are bad at giving information. The British Medical Journal, 292, 1576-1578.
- Maguire, G.P. & Rutter, D.R. (1976). History-taking for medical students: I Deficiencies in performance. The Lancet, 2, 556-558.
- Malpiede, D.M., Leff, M.G., Wilson, K. M. & Moore, V.M. (1982). Assessing interaction between medical trainees and parents of pediatric patients. Journal of Medical Education, 57, 696-700.

- Meen, D.J. (1986). Enhancement of communicated empathy through modeling and corrective feedback in-process of dyadic practice. Unpublished doctoral dissertation, University of Manitoba, Winnipeg.
- Messenger, O.J. (1989). The patient's perception of care: A factor in medical litigation. Canadian Family Physician, 35, 133-135.
- Mills, D.H. & Zytowski, D.G. (1967). Helping relationship: A structural analysis. Journal of Counseling Psychology, 14, 193-197.
- Noren, J. Frazier, T., Altman, I. & DeLozier, J. (1980). Ambulatory medical care: A comparison of internists and family-general practitioners. The New England Journal of Medicine, 302, 11-16.
- Norman, G.R., Neufeld, V.R., Walsh, A., Woodward, C.A. & McConvey (1985). Measuring physicians' performances by using simulated patients. Journal of Medical Education, 60, 925-934.
- Norman, G.R., Tugwell, P. & Feightner, J. W. (1982). A comparison of resident performance on real and simulated patients. Journal of Medical Education, 57, 708-715.
- Poole. A.D. & Sanson-Fisher, R.W. (1979). Understanding the patient: A neglected aspect of medical education. Social Science & Medicine 13A, 37-43.

- Ponterotto, J.G. & Furlong M.J. (1985). Evaluating counselor effectiveness: A critical review of rating scale instruments. Journal of Counseling Psychology, 32, 597-616.
- Rogers, C. (1957). Client-centred theory. Journal of Counseling Psychology, 4, 115-120.
- Sanson-Fisher, R.W. & Maguire, P. (1980). Should skills in communicating with patients be taught in medical school? The Lancet, 2, 523-526.
- Sanson-Fisher, R.W. & Poole, A.D. (1979). Teaching medical students communication skills: An experimental appraisal of the short and long term benefits. In D.J. Osborne, M.M. Gruneberg & J.R. Eiser (Eds.), Research in psychology and medicine (pp 288-294). London: Academic Press.
- Sanson-Fisher, R.W. & Poole, A.D. (1980). Simulated patients and the assessment of medical students' interpersonal skills. Medical Education, 14, 249-253.
- Sideris, D.A., Tsouna-Hadjis, P, Toumanidis, S.T., Vardas, P.E. & Mouloupoulos, S.D. (1986). Medical Education, 20, 307-313.
- Simek-Downing, L. & Quirk, M. (1985). Videotape analyses of medical students' interviewing skills. Family Medicine, 17, 57-60.

- Stewart, M., Brown, J.B. & Weston, W.W. (1989). Patient-centred interviewing part III: Five provocative questions. Canadian Family Physician, 35, 159-161.
- Stewart, M.A. (1984). What is a successful doctor patient interview? A study of interactions and outcomes. Social Science and Medicine, 19, 167-175.
- Stiles, W.B., Putman, S.M., Wolf, M.H. & James, S.A. (1979). Interaction exchange structure and patient satisfaction with medical interviews. Medical Care, 17, 667-679.
- Stillman, P.L., Burpeau-DiGregorio, M.Y., Nicholson, G.I., Sabers, D.L. & Stillman, A.E. (1983). Six years of experience using patient instructors to teach interviewing skills. Journal of Medical Education, 58, 941-945.
- Stillman, P.L., Egan, M.B. & Swanson, D.B. (1987). A diagnostic fourth-year performance-based assessment. Archives of Internal Medicine, 147, 1981-1985.
- Stillman, P.L., Swanson, D.B., Smees, S., Stillman, A.E., Ebert, T.H. & Emmel, V.S., Caslowitz, J., Greene, H.L., Hamolsky, M., Hatem, C., Levenson, D.J., Levin, R., Levinson, G., Ley, B., Morgan, J., Parrino, T., Robinson, S. & Williams, J. (1986). Assessing clinical skills of residents with standardized patients. Annals of Internal Medicine, 105, 762-771.

- Stone, G.C. (1979). Patient compliance and the role of the expert. Journal of Social Issues, 35, 34-59.
- Vaccarino, J.M. (1977). Malpractice: The problem in perspective. Journal of the American Medical Association, 238, 861-863.
- Waitzkin, H. (1984). Doctor-patient communication: Clinical implications of social scientific research. Journal of the American Medical Association, 252, 2441-2446.
- Ward, N.G. & Stein, L. (1975). Reducing emotional distance: A new method to teach interviewing skills. Journal of Medical Education, 50, 605-614.
- Williams, R.G., Barrows, H.S., Vu, N.V., Verhulst, S.J., Colliver, J.A., Marcy, M. & Steward, D. (1987). Direct, standardized assessment of clinical competence. Medical Education, 21, 482-489.
- Wolraich, M.L., Albanese, M., Reiter-Thayer, S. & Barratt, W. (1982). Factors affecting physician communication and parent-physician dialogues. Journal of Medical Education, 57, 621-625.

APPENDIX A

INTERPERSONAL SKILLS RATING SCALE

Station Name: _____

Student No: _____

Simulator's Name: _____

On the next page are listed a number of statements that describe a variety of ways that one person could feel or behave in relation to another person. Please consider each statement with respect to whether you think it is true or not true (false) about your relationship with the student doctor in the interview you have just had. Circle a number in the right margin according to how strongly you feel the statement is true or not true (false) to stand for the following answers:

+3: Yes, I strongly feel that it is true.

-1: No, I feel that it is probably untrue, or more untrue than true (probably false).

+2: Yes, I feel it is true.

-2: No, I feel it is not true (false).

+1: Yes, I feel that it is probably true, or more true than untrue.

-3: No, I strongly feel that it is not true (strongly false).

PLEASE MARK EVERY ONE.

		Strongly True	True	Probably True	Probably False	False	Strongly False
1.	The student doctor wanted to understand how I saw things.	+3	+2	+1	-1	-2	-3
2.	The student doctor looked at what I did from his/her own point of view.	+3	+2	+1	-1	-2	-3
3.	The student doctor usually sensed or realized what I was feeling.	+3	+2	+1	-1	-2	-3
4.	The student doctor just took no notice of some things that I thought or felt.	+3	+2	+1	-1	-2	-3
5.	The student doctor understood me.	+3	+2	+1	-1	-2	-3
6.	The student doctor's response to me was usually so fixed and automatic that I didn't really get through to him/her.	+3	+2	+1	-1	-2	-3
7.	I was able to explain my problem to the student doctor as fully as I needed to.	+3	+2	+1	-1	-2	-3
8.	The student doctor explained things so that now I know what is wrong with me.	+3	+2	+1	-1	-2	-3
9.	The student doctor explained what treatment, tests or other follow up is going to happen.	+3	+2	+1	-1	-2	-3
10.	The student doctor gave me the opportunity to express my feelings or ideas in planning treatment, tests or follow up.	+3	+2	+1	-1	-2	-3
11.	The student doctor often spoke in language I didn't understand.	+3	+2	+1	-1	-2	-3
12.	The student doctor gave me the opportunity to ask questions.	+3	+2	+1	-1	-2	-3
13.	The student doctor was not as thorough as he/she should have been.	+3	+2	+1	-1	-2	-3
14.	I feel satisfied with the medical care that I received.	+3	+2	+1	-1	-2	-3

APPENDIX B

Examinee's I.D.#: _____ Evaluator: _____

Date: _____ Time: _____

PATIENT RATING FORM: PERCEPTIONS OF PROFESSIONAL SERVICE

	EXCELLENT	SATISFACTORY	POOR			
A. COMMUNICATION						
1. <u>Clarity of communication</u> (organization, non-technical language, understandable, purposefulness of questions) <u>Comments:</u>	5	4	3	2	1	
<hr/>						
						Not Done
2. <u>Thoroughness of the explanation of medical problem and management options.</u> <u>Comments:</u>	5	4	3	2	1	0
<hr/>						
B. PROFESSIONAL SERVICE						
1. <u>Professional manner</u> (thoroughness, carefulness, competence, concentration, decisiveness) <u>Comments:</u>	5	4	3	2	1	
<hr/>						
2. <u>Personal Manner</u> (courtesy, respect, sensitivity, earnestness, mannerisms, appearance) <u>Comments:</u>	5	4	3	2	1	
<hr/>						
C. OVERALL						
1. <u>Overall service provided</u> (patient satisfaction, trust, anticipated compliance with treatment plans) <u>Comments:</u>	5	4	3	2	1	
<hr/>						

**APPENDIX C**

THE UNIVERSITY OF MANITOBA

FACULTY OF MEDICINE
Medical EducationS204 -- 750 Bannatyne Avenue
Winnipeg, Manitoba
Canada R3E 0W3
(204) 788-6568

November 27, 1987

To All 4th Year Clinical Clerks

I would like to advise all 4th year clerks about some research I will be undertaking in connection with the Comprehensive Clinical Examination beginning November 30.

I will be investigating certain aspects of the interpersonal skills of the students in their interactions with the simulated patients. I will be using the results for my thesis for a Master of Education degree. The results being analyzed will be for the students as a group and no results of any individual student will be used nor will any individual student be identified.

If any student wishes further information about this, please ask to speak to me by contacting Vera Siemens in the S204 office. Thank you.

Sincerely,

Gail Schnabl, M.S.W.

APPENDIX D

RELATIONSHIPS ACROSS STATIONS BETWEEN IPS TOTAL SCORES AND
COMMUNICATION AND PROFESSIONAL SERVICE SCORES,
BY CORRELATION COEFFICIENTS

STATION	COMMUNICATION	PROFESSIONAL SERVICE
Towns (#1)	.85 (.61) ¹	.86 (.65)
Marchetti (#2)	.80 (.74)	.83 (.73)
Lawrence (#3)*	.73 (.75)	.70 (.72)
Taylor (#4)	.59 (.58)	.85 (.73)
Larsen (#8)	.84	.80
Juranek (#10)	.76 (.76)	.71 (.68)
Higgins (#11)	.58 (.61)	.69 (.74)***
Strawson (#12)	.34 (.51)**	-.06 (.33)
Simcoe (#13)	.84 (.85)	.85 (.86)
McCrae (#14)	.70 (.79)**	.68 (.78)***
Hansen (#15)	.72 (.65)	.70 (.58)
Kenny (#16)	.78	.70
Edwards (#17)	.79	.63
Furlow (#18)	.72 (.70)	.72 (.71)
Jones (#19)	.49 (.54)	.48 (.56)
Hutton (#20)	.73 (.59)	.64 (.63)

- ¹ Correlations using modified scores appear in brackets.
* IPS Total scores did not require modification in these stations.
** Communication scores did not require modification in these stations.
*** Professional Service scores did not require modification in these stations.

APPENDIX E**INSTRUCTIONS TO STUDENTS**

Please complete the attached Interpersonal Rating Scale as you believe the simulated patient in the last station of the exam* has assessed you.

STUDENT # _____

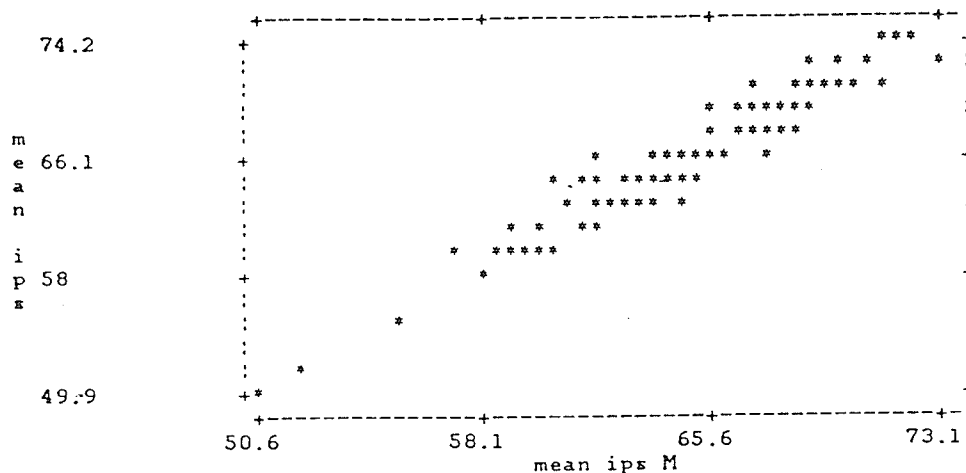
LAST STATION _____

* If your last station was "Mandy Edwards", please use the previous station for your assessment.

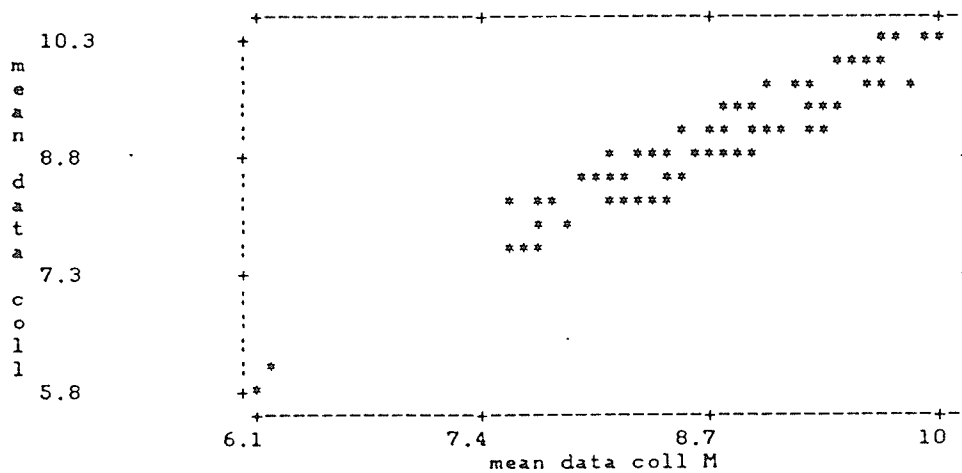
APPENDIX F

RELATIONSHIPS BETWEEN ORIGINAL AND MODIFIED SCORES

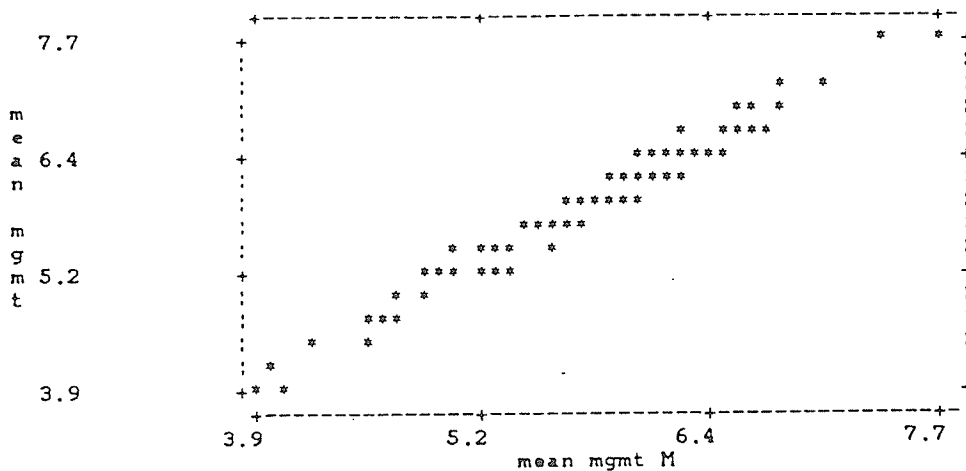
MEAN INTERPERSONAL SKILL SCORES, ORIGINAL AND MODIFIED (M)



MEAN DATA COLLECTION SCORES, ORIGINAL AND MODIFIED (M)



MEAN MANAGEMENT SCORES, ORIGINAL AND MODIFIED (M)



ETHICAL APPROVAL OF RESEARCH AND EXPERIMENT DEVELOPMENT PROJECTS
INVOLVING HUMAN SUBJECTS

This form is to be completed in accordance with the Faculty of Education policy on ethical review. This policy requires that Committee members take into account the relevant standards of the discipline concerned as well as, where appropriate, the standards specified by certain external funding bodies.

Project identification

(to be filled in by investigator)

Investigator(s) Gail K. Schnabl

Title The Interpersonal Skill Competency of
Graduating Medical Students

If applicant is a student, name the faculty member supervising the proposed research

Dr. Clare Pangman

This is to certify that the Review Committee has examined the research and experimental development project indicated above and concludes that the research meets the appropriate standards of ethical conduct in research with human subjects.

Date: Nov. 9 - 1987 Signature of Chairperson _____