THE RELATIONSHIP BETWEEN GRADES, CLINICAL COMPETENCE, AND ENTRY LEVELS AMONG FIRST-YEAR OSTEOPATHIC MEDICAL STUDENTS

Dissertation for the Degree of Ph. D.

MICHIGAN STATE UNIVERSITY

RONALD JAMES MARKERT

1976

This is to certify that the

thesis entitled

The Relationship Between Grades, Clinical Competence, and Entry Levels Among First-Year Osteopathic Medical Students

presented by

Ronald James Markert

has been accepted towards fulfillment of the requirements for

Ph.D. degree in Education

Don E. Hamachek

Major professor

Date May 6, 1976

O-7639

Z-251

T
COM

Pri

research ex

clinical co

physician h

his medical

internships

study exami

competence

laboratory

the relatio

between ent

The results

are (1) an

of a compet

medical edu

grading pro

medical edu

for screeni

procedures.

ABSTRACT

THE RELATIONSHIP BETWEEN GRADES, CLINICAL COMPETENCE, AND ENTRY LEVELS AMONG FIRST-YEAR OSTEOPATHIC MEDICAL STUDENTS

By

Ronald James Markert

Prior to the current study, medical education research examining the relationship between grades and clinical competence was undertaken after the student-physician had completed the classroom-laboratory portion of his medical training. Data typically were gathered during internships, residencies, or general practice. The current study examines the relationship between grades and clinical competence while the student-physicians are in the classroom-laboratory portion of their medical training. In addition, the relationships between entry levels and grades and between entry levels and clinical competence are examined. The results of the analysis of these three relationships are (1) an important preliminary step to the establishment of a competency-based medical education model, (2) an aid to medical educators intent on evaluating the validity of their grading procedures, and (3) a source of information for medical educators interested in establishing entry levels for screening, placement, and individualized instruction procedures.

Eig

for Michiga
Medicine's
students we
week instru
(i.e., neur
mid-term an
two TV case
Problem) wi
During the
complete ne
inations wi
with the si
to perform
ical compet
to the Neur
of the stud
skill, and
the first w
of cognitiv

Mul
grades were
at less tha
regression
the depende

Ronald James Markert

Eighty-seven osteopathic medical students enrolled for Michigan State University's College of Osteopathic Medicine's Neuromuscular Instructional System for first-year students were the subjects for the study. During the ten-week instructional system, data were collected for grades (i.e., neuroanatomy mid-term and final tests, neurosciences mid-term and final tests, a neuroanatomy practical test, two TV case evaluations, and a televised Patient Management Problem) with each measure being appropriately weighted. During the instructional system each student performed two complete neurological evaluation history and physical examinations with simulated patients. The examinations together with the simulated patient ratings of the student's ability to perform the examinations provided the data for the clinical competence variable. Entry levels were gathered prior to the Neuromuscular Instructional System (i.e., the rating of the student's history-taking skill, physical examination skill, and doctor-patient relationship skill) and during the first week of the instructional system (i.e., the test of cognitive knowledge in neurology).

Multivariate regression analysis revealed that grades were related to clinical competence (significant at less than .0001). Furthermore, the most parsimonious regression equation for prediction purposes included only one dependent variable--the scores on the two complete

neurologi

with the

equation.

M

that entr

'signific

parsimoni

included

skill an

knowledge

were the

regressi

on the t

physical

regressi

of the r

to be re

Furtherm

predicti

the test

taking s

relation

of their

neurological evaluation history and physical examinations--with the patient ratings dropping out of the regression equation.

Multivariate multiple regression analysis revealed that entry levels were related to clinical competence (significant at less than .0270). Furthermore, the most parsimonious regression equation for prediction purposes included only two independent variables--history-taking skill and physical examination skill. The test of cognitive knowledge in neurology and doctor-patient relationship skill were the two independent variables dropping out of the regression equation. One dependent variable--the scores on the two complete neurological evaluation history and physical examinations--remained in the most parsimonious regression equation with the patient ratings dropping out of the regression equation.

Multiple regression analysis revealed entry levels to be related to grades (significant at less than .0044). Furthermore, the most parsimonious regression equation for prediction purposes included only one independent variable--the test of cognitive knowledge in neurology. History-taking skill, physical examination skill, and doctor-patient relationship skill dropped out of the regression equation.

The research results were discussed in the context of their implications for future research and practice in

edical ed
n particu
First, the
clinical c
nstructio
research i
competence
was recomm
nto the r
clinical c
hat the "
systematic
nto entry
placement,
specialize
eplicatio
of specifi
sion equat
greater va
ical educa
and validi
This broad
criterion-
the compet
utilizatio

medical education in general and competency-based instruction in particular. A number of considerations were presented. First, the feasibility of using both grading devices and clinical competence measures within the competency-based instruction framework was suggested. Second, replication research investigating the relationship between clinical competence and grades in other areas of medical education was recommended. Third, the value of further investigation into the relationship between clinical skills measures and clinical competence was pointed out. Fourth, it was urged that the "intervening experiences hypothesis" be studied systematically. Fifth, it was recommended that research into entry level variables appropriate for screening, placement, and individualized instruction within other specialized medical education courses be conducted. Sixth, replication research designed to investigate the elimination of specific measures from various most parsimonious regression equations was suggested. Seventh, it was urged that a greater variety of grading devices be introduced into medical education courses and that improving their reliability and validity be the concern of measurement specialists. This broader approach to grading could be a step toward criterion-referenced testing, an essential component of the competency-based instructional model. Eighth, the utilization of physician's assistants in the Neuromuscular

Instructio

insights o

patient p

were desc

patient r

settings

Instructional System was discussed. Ninth, the educational insights gained through the implementation of the simulated patient program for the Neuromuscular Instructional System were described. Tenth, the refinement of the simulated patient rating form and its potential use in other clinical settings were discussed.

T

CON

in

Dep

THE RELATIONSHIP BETWEEN GRADES, CLINICAL COMPETENCE, AND ENTRY LEVELS AMONG FIRST-YEAR OSTEOPATHIC MEDICAL STUDENTS

By

Ronald James Markert

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Counseling, Personnel Services
and Educational Psychology

1976

I w

people whos
me to compl
doctoral re

Fir
doctoral gu
contributio
dissertatio
dissertatio
tion during
been a teac
His teachin
professiona
become a ha
allowed me
of his Neuro
willingness
task. Dr.
role model
observing hi
fitting from

ACKNOWLEDGMENTS

I would like to express my appreciation to those people whose concern and effort were instrumental in helping me to complete my doctoral program in general and my doctoral research in particular.

First, I would like to extend my gratitude to the doctoral guidance committee; each member made a significant contribution to my doctoral program. Dr. Fred C. Tinning, dissertation director, not only gave superb guidance to the dissertation but also has been my mentor in medical education during the last two years. Dr. Don E. Hamachek has been a teacher, academic advisor, counselor, and friend. His teaching and educational insights have made me a better professional; his concern and friendship have helped me become a happier person. Dr. Lawrence Jacobson kindly allowed me to collect my research data within the context of his Neuromuscular Instructional System. His interest and willingness to help made the research project an enjoyable task. Dr. Edward L. Smith has been for me an outstanding role model of a dedicated, thorough researcher. Through observing him while conducting his own research and benefitting from the assistance he gave me in conducting my own

research,

developmen

Da

research d

I am most

the disser

shared dur

thank Davi

assistants

which made

I w

Gordon, who

who provide

patients, a

and Patholo

of his neur

Fou

I am apprec

Linda Phill

A s

patients wh

the College

were the su

research, Dr. Smith has greatly contributed to my development as an educational researcher.

David West was always willing to assist with my research design, measurement, and statistical problems. I am most grateful for both his assistance in completing the dissertation and the many learning experiences we have shared during the last three years. I would also like to thank David Harrison and Kenneth Klegon, who, as physician's assistants, provided instructional and research functions which made the research project possible.

I would like to recognize the assistance of Eric Gordon, who helped me with data processing, Holly Holdman, who provided insights into the training of simulated patients, and Dr. Thomas W. Jenkins, Professor of Anatomy and Pathology, who patiently explained to me the content of his neuroanatomy tests.

Four people provided vital clerical support. I am appreciative of the help given by Karen Ammarman, Linda Phillipich, Brenda Spraggins, and Grace Rutherford.

A special thanks is extended to the simulated patients who participated in the research project and to the College of Osteopathic Medicine's Class of 1977, who were the subjects for the study.

I am

Medicine and

with facilit

Fina

and interest

been the sou

both persona

I am grateful to the College of Osteopathic Medicine and Dean Myron Magen for supporting my research with facilities and monies.

Finally, I wish to acknowledge the enduring love and interest of my parents, John and Elsie. They have been the source from which I have drawn strength for both personal and professional growth.

TABLE OF CONTENTS

Chapter

Hypo

Summ
Addi

V. DISCU
THEIR
AND P

Rese
fo

Prac
St

Summ

Chapter | Page

Appendix

Table

1.

2.

3.

4.

5.

6.

7.

8.

9.

10.

11.

LIST OF TABLES

Table

Thi

between aca

entry level

and variati

in the lite

is, in acco

study.

Academic Pe

Con

academic pe

average (GI

sometimes

has faded

In all stud

achievemen

Ope

academic pe

in the Neur

weighted T-

see Chapter

DEFINITION OF TERMS

This study is concerned with the relationship between academic performance, clinical competence, and entry levels. For purposes of lucidity, these three terms and variations of them will be defined conceptually as used in the literature in general and then operationally, that is, in accordance with how they are used in the current study.

Academic Performance

Conceptual definition.--In the studies cited, academic performance usually refers to either grade-point average (GPA) or class rank. However, earlier studies sometimes used a measure of academic performance which has faded from contemporary use (e.g., a percentage grade). In all studies cited, academic performance refers to achievement in grades.

Operational definition.--In the current study, academic performance is determined by weighting the tests in the Neuromuscular Instructional System to produce a final weighted T-score for each student. (For further description see Chapter III.)

Clinical Co

Co

the Commit

of Medical

ability an

diagnosis,

theoretica

includes s

informatio

symptoms a

reasonable

of patient

Op

clinical co

score on t

physical e

from the r

patients.

Entry Leve

Co

the aptitu

student pos

instruction

Ope

entry level

Clinical Competence

Conceptual definition.--According to the Report of the Committee on Goals and Priorities of the National Board of Medical Examiners (1973), clinical competence is "the ability and/or qualities requisite for patient care, diagnosis, treatment, and management as distinguished from theoretical or experimental knowledge. Clinical competence includes such elements as skill in obtaining pertinent information from a patient, ability to detect and interpret symptoms and abnormal signs, acumen in arriving at a reasonable diagnosis, and judgment in the management of patients" (p. 86).

Operational definition.--In the current study, clinical competence is composed of (1) the student's mean score on two complete neurological evaluation history and physical examinations and (2) the student's mean score from the ratings of his performance by his two simulated patients. (For further description see Chapter III.)

Entry Level(s)

Conceptual definition.--Entry level(s) refer(s) to the aptitude(s) related to a course of instruction which a student possesses prior to beginning that course of instruction.

Operational definition.--In the current study, four entry levels for each student are utilized: (1) score on a test of cog

history-tak

course, (3

during a ph

patient rel

diagnosis c

Career Perf

Con

encompasses

efforts. I

involving e

for career

In the medi

times the s

Professiona

the synonym

Ope

is not used

the relevan

test of cognitive knowledge in neurology, (2) score on history-taking skill as measured during a physical diagnosis course, (3) score on physical examination skill as measured during a physical diagnosis course, and (4) score on doctor-patient relationship skill as measured during a physical diagnosis course. (For further description see Chapter III.)

Career Performance

Conceptual definition.--Career performance encompasses the criteria for evaluation of a subject's efforts. For example, Taylor (1963) reports a study involving engineers and physicists where the criteria for career performance were creativity and productivity. In the medical studies cited, clinical competence is sometimes the sole or principal component of career performance. Professional performance and occupational success are among the synonyms used in this text for career performance.

Operational definition.--The term career performance is not used in the current study. Clinical competence is the relevant equivalent for the study herein conducted.

N

Th

has been a

last decad

Education,

instructio

medical ed

tives (His

and Kofman

simulated

maximal pa

and evalua

experience

the challe

numbers th

efficacy o

techniques

training o

CHAPTER I

INTRODUCTION TO THE STUDY

Need for the Competency-Based Instructional Model in Medical Education

The public's demand for better health care services has been a persistent theme for medical education in the last decade (Coggeshall, 1965; Carnegie Commission on Higher Education, 1970; Millis, 1971). Consequently, curricular, instructional, and technological innovations have pervaded medical education in recent years (e.g., behavioral objectives (Hiss and Peirce, 1974), simulation devices (Penta and Kofman, 1973), simulated patients (Thinning, 1973), simulated patient management problems (Goran et al., 1973), maximal participation of students in patient examinations and evaluations (Talalla et al., 1974), early interviewing experience (Goroll et al, 1974)) with the goal of meeting the challenge of producing competent physicians in greater numbers than ever before at the lowest cost possible. The efficacy of such *micro* approaches (i.e., instructional techniques applicable within units of instruction) to training competent physicians is well-documented.

At

approach to

fessional

training.

satisfacto

i.e., the

Examiners,

Examiners

licensing

examinatio

Priorities

1973) has

Board exam

profession

tions in s

have admin

principles

examinatio

Wh

is evident

both the

clinically

classroom

of medical

education

At the other end of the spectrum, the *macro* approach to assuring competence attempts to assess professional competence as an end product of medical school training. As a licensing requirement, each state requires satisfactory completion of a certification examination (i.e., the examinations of the National Board of Medical Examiners, the examinations of the National Board of Examiners for Osteopathic Physicians and Surgeons, Federal Licensing Examinations, or individual state medical board examinations). The Report of the Committee on Goals and Priorities of the National Board of Medical Examiners (1973) has proposed a number of alterations in the National Board examinations designed to assure better measurement of professional competence. Specialty certification examinations in such areas as orthopedics and internal medicine have admirably applied sound learning and measurement principles in an effort to develop competency-based examinations.

While the heightened emphasis on clinical competence is evident from the work of medical educators who pursue both the *micro* and *macro* approaches, the quest to train clinically competent physicians must also be focused on the classroom and laboratory courses which are the foundation of medical education. There is a striking need in medical education for the development of a competency-based instruction

are clinica

accordance

a major ste

greater cli

consequentl

Whi

origins, an

II, it woul

briefly the

instruction

model speci

criterion f

mining whet

tion are in

learner to

competence)

objectives,

instructior

accountable

Pre

bound model

need to be

science and

what degree

instructional model which can be applied to courses that are clinically oriented. Such courses, developed in accordance with an empirically sound model, would be a major step toward assuring improved physician training, greater clinical competence among physicians, and consequently better health care for the public.

While competency-based instruction--its definition, origins, and development--is discussed in detail in Chapter II, it would be advantageous at this point to mention briefly the core characteristics of the competency-based instructional model. The competency-based instructional model specifies at the outset and in performance terms the criterion for competence and indicates the means for determining whether the criterion is reached. Modes of instruction are implemented which will be optimal in aiding the learner to achieve the objectives (i.e., demonstrate competence). In the competency-based instructional model, objectives, criterion, means of assessment, and modes of instruction are publicly shared, and the learner is held accountable for meeting the criterion.

Preliminary to the development of such a competency-bound model for medical education, a variety of questions need to be examined--e.g., (1) How should the relevant basic science and clinical science material be integrated; (2) To what degree can medical education be self-paced; (3) How can the compon

optimally

investigat

model invo

this respe

1. To

   co

2. To

   kn

   ti

   pr

Wi

ure from

performanc

relation b

reveal, cl

of career

finding, s

pening exp

effect tha

attitudes,

and the ti

the components of a competency-based instruction model be optimally utilized in medical education? Another area of investigation preliminary to developing a competency-based model involves student aptitude and achievement, and in this respect two questions seem basic:

1. To what degree do grades predict clinical competence?
2. To what degree do entry levels (i.e., cognitive knowledge, history-taking skill, physical examination skill, and doctor-patient relationship skill) predict clinical competence?

Grades as Predictors of Clinical Competence

Wingard and Williamson (1973) reviewed the literature from 1955 to 1972 on grades as predictors of career performance among physicians and found little or no correlation between the two variables. (As Chapter II will reveal, clinical competence often was not the criterion of career performance.) One possible explanation for this finding, suggest Wingard and Williamson, is that the intervening experiences--e.g., internships, residencies, the effect that the demands of medical practice have on habits, attitudes, and interpersonal skills--between medical school and the time when the measure(s) of career performance is (are) gathe

between aca

Evi

correlation

study in wh

be correla

completed

suggest th

school and

distort th

petence.

irrelevant

have to re

a nonsigni

study, the

vening ex

changing

competenc

T

relations

in-traini

developme

should be

(are) gathered may significantly distort any relationship between academic performance and professional performance.

Evidence in regard to this explanation of the low correlation between the two variables could come from a study in which academic performance (i.e., grades) would be correlated with clinical competence before the subjects completed medical school. A significant correlation would suggest that the intervening experiences between medical school and the gathering of data on career performance do distort the relationship between grades and clinical competence. Thus, those who claim that grading procedures are irrelevant to professional performance for physicians may have to reconsider their position. On the other hand, if a nonsignificant correlation is found in the above-mentioned study, then the argument stressing the importance of intervening experiences will be weakened, and advocates of changing grading procedures to better reflect clinical competence will have support for their position.

To date there are no studies investigating the relationship between these two variables among physicians-in-training. Studies of this nature are relevant to the development of a competency-based instructional model and should be promoted by medical educators.

The

instruction

an understa

levels of s

competence.

scores on s

(2) history

examination

significant

petence for

levels migh

instruction

student cou

achieve the

proceeding

examining

clinical c

systems so

level prer

medical ed

Entry Levels as Predictors of Clinical Competence

The development of a model for competency-based instruction in medical education would be facilitated by an understanding of the relationship between the entry levels of students and their achievement of clinical competence. In other words, if a student's entry level scores on such variables as (1) cognitive medical knowledge, (2) history-taking skill, (3) skill in performing a physical examination, and (4) ability to relate to a patient are significantly correlated with achievement of clinical competence for a particular body system, then required entry levels might be prescribed before a student could begin instruction. As part of the competency-based model, a student could be required to undergo remediation and to achieve the prescribed entry levels before beginning or proceeding further in an instructional unit. Research examining this question of entry levels as predictors of clinical competence needs to be done within various body systems so as to provide a data base for developing entry level prerequisites as part of a competency-based model for medical education.

Th

the relati

and (2) en

two questi

their impl

model of m

study. Al

performanc

revealed l

regard to

career per

directly re

after the c

Inv

levels and

educators w

tional mode

vidualized

principally

medical edu

ment of a

While the c

competency-

to medical

Purpose

This study is primarily interested in investigating the relationship between (1) grades and clinical competence and (2) entry levels and clinical competence. While these two questions are of general interest to medical educators, their implications for the development of a competency-based model of medical instruction has stimulated the present study. Although correlational studies relating academic performance (typically grades) to career performance have revealed little or no relationship, these studies suffer in regard to relevance to a competency-based model in that the career performance measure is often based on criteria not directly related to clinical competence and is gathered after the completion of medical training.

Investigation of the relationship between entry levels and clinical competence is important to medical educators who want to establish a competency-based instructional model incorporating screening, placement, and individualized instruction procedures. Thus, this study focuses principally on these two questions of general interest to medical education and of specific importance to the development of a competency-based medical instruction model. While the current study is theoretically linked to competency-based instruction, the findings may be valuable to medical educators in ways unanticipated by the researcher.

In a

indings to

the current

ation with

State Unive

researcher

physician's

competence

patients ca

setting. T

Chapter V.

Co

University

mittee on

improved e

further re

ratings of

physical e

relationsh

the study

recommenda

instructio

T

Chapter w

The study

In addition to the application of the research findings to the development of a competency-based model, the current study has potential for more immediate application within the College of Osteopathic Medicine, Michigan State University. In conjunction with the study, the researcher and his colleagues implemented a novel use for physician's assistants (PAs) as facilitators of clinical competence and further explored means by which simulated patients can be used effectively in a clinical learning setting. These two activities will be described in Chapter V.

Consistent with the report of the Michigan State University College of Osteopathic Medicine's Ad Hoc Committee on Evaluation Procedures (1975), which called for improved evaluation of clinical experiences, the researcher further refined an earlier instrument used to obtain patient ratings of a student's ability to perform a history and physical examination. Finally, in its investigation of the relationship between entry levels and clinical competence, the study provides impetus to the ad hoc committee's recommendation for screening, placement, and individualized instruction.

The review of literature to follow in the next chapter will make clear the uniqueness of the current study. The study herein researched and reported is, to the researcher's

examines th

and clinica

clinical co

classroom-

during the

training (r

a whole).

The

1. For
a n
com
eva
not

2. For
a n
com
eva
not

To

equations u

clinical co

following n

3. For
a n
are

The results

hypothesis r

researcher's knowledge, the first investigation which examines the relationship between academic performance and clinical competence and between entry levels and clinical competence *while the student-physician is in the classroom-laboratory portion of his medical training and during the course of a segment of that classroom-laboratory training* (rather than the student-physician's training as a whole).

Null Hypotheses

The null hypotheses of principal interest are:

1. *For first-year osteopathic medical students in a neuromuscular instructional system, clinical competence in performing a complete neurological evaluation history and physical examination is not related to grades.*

2. *For first-year osteopathic medical students in a neuromuscular instructional system, clinical competence in performing a complete neurological evaluation history and physical examination is not related to entry levels.*

To complete the testing of all meaningful regression equations utilizing the three variables of interest--clinical competence, grades, and entry levels, the following null hypothesis will be investigated.

3. *For first-year osteopathic medical students in a neuromuscular instructional system, grades are not related to entry levels.*

The results of the analysis dealing with this third hypothesis may be of interest to medical educators who utilize mor

to introduc

and individ

C

Mic

tional Syst

is the cont

operational

Grades-
mea
Ins

Clinica
neu
ina
rat
his

Entry l
neu
(3)
rat

Thi

the need fo

medical edu

of the conce

origins and

literature c

and career p

utilize more traditional grading devices and who want to introduce entry level measures for screening, placement, and individualized instruction purposes.

Operational Definitions of the Variables

Michigan State University's Neuromuscular Instructional System for first-year osteopathic medical students is the context for data collection. The variables are operationally defined as follows:

Grades--overall weighted T-score derived from eight measures used for grading during the Neuromuscular Instructional System.

Clinical competence--(1) mean T-score for two complete neurological evaluation history and physical examinations and (2) mean T-score for two patient ratings of two complete neurological evaluation history and physical examinations.

Entry levels--(1) test of cognitive knowledge in neurology, (2) rating of history-taking skill, (3) rating of physical examination skill, (4) rating of doctor-patient relationship skill.

Overview

This chapter has introduced the study focusing on the need for the competency-based instructional model in medical education and the purpose of the study. In Chapter II the concept of competency-based instruction and its origins and development will be reviewed as will the literature on the relationship between academic performance and career performance and the relationship between entry

levels and
and its cha
generalizat
MSU-COM's N
research de
outlined, t
hypotheses
dures descr
the analysi
most parsin
tion which
is also rep
results of
future rese

levels and clinical competence. In Chapter III the sample and its characteristics are discussed, the threats to generalization detailed, the instructional context (i.e., MSU-COM's Neuromuscular Instructional System) of the research described, the various measures used in the study outlined, the calendar for data collection and the research hypotheses presented, and the statistical analytic procedures described. Chapter IV reports for each hypothesis the analysis of the data, the status of the hypothesis, the most parsimonious regression equation, and the interpretation which can be drawn. Additional analysis of the data is also reported in Chapter IV. Chapter V discusses the results of the study and presents the implications for future research and practice.

Thi

by focusing

based instr

origins, ar

dealing wit

and career

been subdi

business,

Third, ent

be examine

examining

assessed j

instructio

concentra

by invest

and preme

to charac

successfu

tion of

elements

CHAPTER II

REVIEW OF LITERATURE

Organization

This chapter places the current study in perspective by focusing on three principal topics. First, competency-based instruction will be discussed--a definition, its origins, and its development. Second, the related research dealing with the relationship between academic performance and career performance will be reviewed. This section has been subdivided into five professional areas--teaching, business, engineering, science research, and medicine. Third, entry level predictors of clinical competence will be examined. In this section, due to a lack of research examining the relationship between relevant entry levels assessed in close proximity to instruction and subsequent instruction (a focus of the current study), the review will concentrate on the two prime entry level variables pursued by investigators--the Medical College Admission Test (MCAT) and premedical grade point average (GPA). Studies seeking to characterize the personality factors of the typical or successful medical student will also be reported in recognition of the potentially valuable use for noncognitive elements as entry levels.

Definitio

E

instructi

outcome b

an approp

at the co

measures

the presc

to permi

the instr

tional u

competen

easily o

refer to

instruct

1.

2.

3.

4.

5.

6.

Competency-Based Instruction

Definition

Before instruction begins, the competency-based instructional model specifies in performance terms the outcome behaviors which are deemed essential for competence. At appropriate junctures during the instructional unit and at the conclusion of instruction, criterion-referenced measures are taken to ascertain how the student compares to the prescribed criterion levels established as satisfactory to permit the student to continue on to the next part of the instructional unit or to receive a pass for the instructional unit. As might be deduced from the above explanation, competency-based instruction is a concept which is not easily operationally defined. Houston and Howsam (1972) refer to the "central characteristics" of competency-based instruction:

1. specification of learner objectives in behavioral terms;
2. specification of the means for determining whether performance meets the indicated criterion levels;
3. provision for one or more modes of instruction pertinent to the objectives, through which the learning activities take place;
4. public sharing of the objectives, criteria, means of assessment, and alternative activities;
5. assessment of the learning experience in terms of competency criteria; and
6. placement on the learner of the accountability for meeting the criteria [pp. 5-6].

T

cornersto
performan
need to b
performan
learner c
objective
the stude
of instru
fication
meets the
the learn
they are
criterion
level of
a specifi
of instru
tion. In
student w
student h

T

test is
are direc
standards
(1971) de

To supplement Houston and Howsam, the two cornerstones of competency-based instruction (i.e., performance objectives and criterion-referenced tests) need to be defined. Houston and Howsam are referring to performance objectives when they mention "specification of learner objectives in behavioral terms." Performance objectives are specific statements of behaviors which the student should be able to demonstrate as the result of instruction. When Houston and Howsam mention "specification of the means for determining whether performance meets the indicated criterion levels" and "assessment of the learning experience in terms of competency criteria," they are referring to criterion-referenced testing. A criterion-referenced test compares a student's attained level of competence for a particular behavior or skill with a specified competence level judged necessary to pass a unit of instruction or to proceed to the next module of instruction. In contrast, a norm-referenced test compares the student with others in his group or class and grades the student by using the normal distribution.

To Glaser and Nitko (1971) a criterion-referenced test is "deliberately constructed to yield measurements that are directly interpretable in terms of specified performance standards . . ." (p. 653). Furthermore, Harris and Stewart (1971) define a criterion-referenced test as one which consist o

population

may be use

of tasks."

Ho

concepts a

competency

vidualized

systems ap

and manage

contributo

approach.

Origins ar
Competency

Co

in teacher

history of

1967 the U

impetus to

issuing a

undergradu

Steffensor

Models pro

National C

the Natior

"consist of a sample of tasks taken from a well-defined population of tasks such that the performance on the sample may be used to generalize to performance on the population of tasks."

Houston and Howsam (1972) describe a number of other concepts and techniques which often are part of a specific competency-based instructional model's commitment to individualized instruction--e.g., modularized packaging, the systems approach, educational technology, and guidance and management support. However, these components are contributory and not essential to the competency-based approach.

Origins and Development of the Competency-Based Instruction Approach

Competency-based instruction has its formal origins in teacher education. Steffenson (1974), in tracing the history of competency-based instruction, reports that in 1967 the U.S. Office of Education's Bureau of Research gave impetus to Competency-Based Teacher Education (CBTE) by issuing a request for proposals to develop comprehensive undergraduate and inservice teacher education programs. Steffenson notes that the Elementary Teacher Education Models project provided the early CBTE literature, and the National Center for Improvement of Educational Systems and the National Teacher Corps were the most prominent Office of Educ

instruc

College

1972, 1

based c

same su

institu

trainin

that th

instruc

physici

cians,

pointed

menting

nized

McMaste

Missour

College

develop

entitle

the fi

of Education programs which explored CBTE as a means of restructuring teacher education.

A survey conducted by the American Association of Colleges for Teacher Education (1972) reported that by fall 1972, 17 states had devised teacher certification procedures based on the competency-based instruction approach. The same survey reported that of the 783 teacher education institutions responding, 125 reported that their teacher training program was competency-based while 366 indicated that their CBTE program was in the developmental stage.

Houston (1974) observed that the competency-based instruction approach is making inroads in the training of physicians, dentists, nurses, pilots, engineers, electricians, plumbers, and computer programmers.

Jason (personal communication, October 9, 1975) pointed out that the progress of medical schools in implementing competency-based programs is very slow but recognized Michigan State University's College of Human Medicine; McMaster University in Hamilton, Ontario; the University of Missouri at Kansas City; and the University of Illinois, College of Medicine, Chicago, as innovators in the development of the competency-based instruction concept.

Chan et al. (1974) have developed a clerkship entitled *Fundamentals of Patient Care* (FPC), which is the first full-time clinical experience for Michigan State

University's s

to orient th

practice and

demonstrate

the program

in eleven ca

1. Acqu
2. Coll
3. Ment
4. Phys
5. Prob
6. Basi
7. Epid
8. Deve
9. Heal
comm
10. Labo
11. Prof

Consistent w

the time fra

basic unit c

time, some r

Eval

instructiona

listed above

of the stude

to-day basi

patients, (

(3) student

problems, a

X-ray unknow

University's College of Human Medicine students. Designed to orient the student into full-time, realistic medical practice and to give the student the opportunity to demonstrate and further develop his clinical skills, the program meticulously assesses student competence in eleven categories:

1. Acquisition and refinement of interviewing skills
2. Collection and organization of historical data
3. Mental status examination
4. Physical examination
5. Problem solving, problem management, and follow-up
6. Basic mechanisms of disease
7. Epidemiology and preventive medicine
8. Developmental variations in problems
9. Health care delivery systems--particularly the community hospital
10. Laboratory skills and knowledge
11. Professional behavior.

Consistent with many competency-based instruction programs, the time frame for completion of FPC is flexible within the basic unit of 12 weeks, that is, some students require less time, some more time, in which to achieve competence.

Evaluation of FPC students is based on specific instructional objectives derived from the eleven categories listed above. Evaluation takes four forms: (1) observation of the student's performance on both a general on-going day-to-day basis and his specific interaction with particular patients, (2) review of the written records of the student, (3) student performance on standardized simulated clinical problems, and (4) the student's analysis of laboratory and x-ray unknowns.

niversi
such app
ng base
utorial
interper
resource
(5) diag
frequent
learning
and facu

need int
(1972) d
general
educatio
Monthly
the medi
effectiv
of the h
review p
the reas
medical

T
Illinois,

Neufeld and Barrows (1974) describe the McMaster University focus on the individual student physician through such approaches as: (1) self-directed learning, (2) learning based on student-identified problems, (3) small group tutorials designed to enhance both problem-solving and interpersonal skills, (4) a wide variety of learning resources including both real and simulated patients, (5) diagnostic or formative evaluations intended to frequently assess and, if desirable, modify the student's learning progress, and (6) integration of knowledge, skill, and faculty during the learning process.

Based on the philosophy that physicians-in-training need intimate contact with humans and their problems, Dimond (1972) describes the activities of the small, comprehensive general medicine team, a vital component in the medical education program at the University of Missouri, Kansas City. Monthly Docent Audits (UMKC uses the term *docent* to refer to the medicine team) are conducted to assess the cost-effectiveness of the medicine team concept. The quality of the health care provided is assessed by extensive peer review procedures, chart reviews, and daily analysis of the reasoning process by means of the problem-oriented medical record.

Tremonti (1974) reports that the University of Illinois, College of Medicine, Chicago, has developed behavior

Phase I-

physical

a proble

approach

elective

nostic e

the spec

the deve

in an ef

For exam

and comp

examinat

a Patien

(ABIM) c

in crite

contain

but also

Currentl

ported th

an "erasu
requires
immediate

behavioral goals for the four phases of its curriculum: Phase I--information gathering, the interview, the physical examination; Phase II--problem-solving through a problem identification, assessment, and management approach; Phase III--clinical experience; and Phase IV--electives. In addition, the college has developed diagnostic examinations for each of the four phases to test the specific objectives established for each phase.

Medical specialty boards have taken the lead in the development of standardized criterion-referenced tests in an effort to promote a competency evaluation approach. For example, orthopedic board examinations are multiphasic and competency-based consisting of a multiple-choice examination, an oral examination and, until recently, a Patient Management Problem (PMP)[1] examination.

Also, the American Board of Internal Medicine (ABIM) certification examination is considered a leader in criterion-referenced test development. The ABIM exams contain the traditional multiple-choice and true-false items but also an increased number and a broader sample of PMP's (currently 15) in recognition of research which has supported the domain-specific and medical case-specific nature

---

[1]A Patient Management Problem (PMP), commonly called an "erasure test," presents a clinical case or problem which requires the examinee to make sequential decisions based on immediate feedback received from the answer sheet.

of clin

al., 19

tion ex

using

better

Nationa

Surgeon

as an c

and pra

perform

pursued

between

studies

recent

relatio

formanc

busines

before

five ar

have be

of clinical competence (McGuire and Page, 1973; Elstein et al., 1973). In summary, medical specialty board certification exams have introduced PMP's and structural oral exams (using such techniques as role-playing and simulation) to better assess clinical competence.

Finally, Part III of the examination given by the National Board of Examiners for Osteopathic Physicians and Surgeons (1974) assesses the candidate's clinical competence as an osteopathic physician and surgeon through both oral and practical tests.

Academic Performance as a Predictor of Career Performance

Research examining the relationship between academic performance and career performance has never been extensively pursued, and, possibly due to the lack of relationship between the two variables found in the limited number of studies completed, even less interest can be detected during recent years. In reviewing studies which have explored the relationship between academic performance and career performance, research dealing with professionals in teaching, business, engineering, and scientific research will be cited before turning to medical professionals. It is in these five areas that most studies examining the present issue have been undertaken. The section which reports research involvin

involvin

dealing

to signi

cases st

indicate

will dis

signific

Teaching

the rela

performa

College

of each

ful grad

responde

Of these

their cl

Kunkel c

between

findings

ological

(2) only

The judg

involving medical professionals includes two studies involving nurses in addition to the relevant research dealing with physicians.

Throughout this section reference will be made to significant and nonsignificant correlations. In all cases statistical significance or nonsignificance is being indicated. The "Conclusion" at the end of this section will discuss the issue of statistical versus practical significance.

Teaching

Kunkel (1917) inaugurated the investigation of the relationship between academic performance and career performance among teachers with his study of Lafayette College graduates between 1876 and 1905. Ten graduates of each class were asked to nominate the five most successful graduates of their class. Only 123 of the 300 judges responded, and among their 301 nominees were 55 teachers. Of these 55, 62 percent were in the upper one-fifth of their class and only 5 percent in the lowest one-fifth. Kunkel concluded that there was a direct relationship between academic performance and teacher success. Kunkel's findings, however, are suspect due to at least two methodological difficulties: (1) "success" was not defined; and (2) only 150 of the 301 nominees were nominated by more than one judge, indicating an idiosyncratic frame of reference by the jud

a time

practic

College

ing. T

teacher

details

upper,

managem

between

ship be

since 4

rating

middle

similar

of elap

finding

objecti

Payne's

nonrand

relativ

signifi

signifi

the judges. In addition, generalizing these results to a time period more than 50 years later is a questionable practice.

Payne (1918) had 144 graduates of Harris Teachers College rated by their principals after one year of teaching. The three criteria of teaching success on which teachers were rated included management, attention to details, and instruction. Graduates were divided into upper, middle, and lower thirds academically. The criteria management and attention to detail yielded no differences between groups; however, Payne claims a positive relationship between academic performance and success in instruction since 40 percent of the upper third received an "excellent" rating from their principals where only 27 percent of the middle third and 17 percent of the lower third received a similar rating. As with the Kunkel study, over 50 years of elapsed time must qualify the generalizability of any findings. In addition, today's educators would raise strong objections to the limited criteria for success (albeit Payne's criteria were considerably better than Kunkel's).

Gambrill (1922) surveyed the 1903 graduates of 11 nonrandomly selected colleges and correlated salary with relative class rank for 65 men and 95 women teachers. A significant .28 correlation for men was found, but no significant correlation for women (.04) was obtained.

She con
academi
century
In addi
and a I
methodo
Finally
for mer
the tim
scales
impress

Univers
tivenes
classro
possibi
attitud
criteri
were es
with th
variabl
group a
means w
extreme
Questio

She concluded that there was little relationship between academic performance and success in teaching. The half century time lapse again presents generalizability problems. In addition, Gambrill had a nonrandom sample of colleges and a low return rate for her questionnaire. These two methodological problems cast doubt on the findings. Finally, the author suggests that the .28 correlation for men might have been an artifact of the practice of the time whereby school systems with the highest salary scales tended to employ beginning teachers with the more impressive academic records.

Stuit (1937) had school superintendents rate University of Nebraska graduates on seven teaching effectiveness criteria (i.e., teaching results, teaching methods, classroom management, student-teacher relations, growing possibilities as a teacher, cooperation and professional attitude, and personal qualities). On the basis of the criteria, four groups--superior, good, average, and poor--were established. The superior group (N = 100) was compared with the poor group (N = 46) using grades as the dependent variable. The superior group averaged 85.0 and the poor group averaged 82.4. The difference between these two means was statistically significant. Although the two extreme groups were significantly different, one must question the practical or meaningful significance of a

differen
Stuit i
middle
it is d
cedure
have yi

graduat
and pup
achieve
academi
signifi
educati

Univers
teachin
of Wisc
score o
with ni
and col
rating
(rangin
with an
icantly
the 17

difference of 2.6 points on a 100-point scale. In addition, Stuit ignores the vast majority of teachers, those in the middle two groups (i.e., good and average). Consequently, it is doubtful that a more sophisticated statistical procedure (e.g., Pearson's r) utilizing all subjects would have yielded significant results.

Jones (1946) sampled 65 University of Wisconsin graduates (1941-1943) and correlated supervisory rating and pupil gain score (i.e., improvement in standardized achievement test scores) with six measures of undergraduate academic performance. Only one of twelve correlations was significant (.40 between supervisory rating and GPA in educational courses).

Lins (1946) studied 58 1943 graduates of the University of Wisconsin who were in their first year of teaching. Three teaching efficiency criteria--University of Wisconsin faculty rating, pupil rating, and pupil gain score on standardized achievement tests--were correlated with nine GPA values derived from different types of courses and college time periods. University of Wisconsin faculty rating correlated significantly with eight of the nine GPAs (ranging from .28 to .33). Pupil rating did not correlate with any of the GPAs. Pupil gain scores correlated significantly with four of the GPAs (ranging from .52 to .56) for the 17 teachers for whom pupil gain scores were available.

Both gr

First,

likely

teache

the co

rating

scores

rather

.05 no

and sa

who be

of the

in the

and st

betwee

ent GP

coeffi

first

teachi

graduat

of sig

Both groups of significant correlations are suspect. First, the University of Wisconsin faculty members were likely familiar with the academic performance of the teachers, thus suggesting criterion contamination with the correlations between University of Wisconsin faculty rating and GPAs. Second, the correlations for pupil gain scores and GPAs included only 17 of the 58 teachers, a rather small sample.

Jepsen (1951) found no relationship (actually a .05 nonsignificant correlation) between academic performance and salary among 160 Fresno State male graduates (1929-1941) who became teachers.

Erickson (1954) had 64 graduates from the 1950 class of the University of Wisconsin's School of Education rated in their second year of teaching by their principal, peers, and students and by independent observers. The correlations between these measures of career performance and 10 different GPAs ranged from -.29 to +.28. Only two correlation coefficients were significant.

Stoelting (1955) correlated grades during the first two years of college with principal ratings of teaching success for 88 of the 1952 University of Wisconsin graduates. A correlation of .385 is reported but no level of significance is given.

graduat
measure
rating
.29 wit
in the

between
all pro
graduat

GPA to
overall
A & M C
of teac
or othe
tions w
correla
matter,
in Engl
to .38.

effecti
were co
college
.29.

Jones (1956) studied 46 randomly selected female graduates of the University of Wisconsin (1951-1953). The measure of career performance used was the principal's rating of the teacher's ability. This rating correlated .29 with grades in professional courses and .33 with grades in the teaching major. Both correlations were significant.

Schick (1957) found a nonsignificant .05 correlation between supervisor's rating of teaching success and GPA in all professional courses for 72 1955 University of Wisconsin graduates who were in their first year of teaching.

Massey and Vineyard (1958) related undergraduate GPA to 14 qualities important to teacher success and an overall rating of teacher success. Sixty-two Panhandle A & M College graduates (1954-1956) in their first year of teaching were rated on the 15 scales by their principal or other appropriate supervisor. Only 4 of the 15 correlations were significant. The range of the 4 significant correlations--undergraduate GPA with (a) mastery of subject matter, (b) character, standards, and ideals, (c) competence in English expression, and (d) general culture--was from .28 to .38.

Cole (1961) had 140 teachers rated on teaching effectiveness by trained observers. When these ratings were correlated with grades during the first two years of college, the correlation coefficient was a nonsignificant .19.

Simun and Asher (1964) conducted a study in which GPA was correlated with school administrator ratings of teaching ability, discipline, preparation of subject matter content, staff cooperation, and tact with students. The sample was 92 percent (N = 111) of the graduates of the Carnegie Institute of Technology, who were rated by administrators during their first year of teaching. Most correlations were nonsignificant. The range of correlations was from -.17 to +.36.

Nair (1973), in his study of 80 teacher trainees, found academic performance not to be a good predictor of teacher effectiveness.

Business

Kunkel (1917), in his study of 1876 to 1905 Lafayette College graduates, had 50 subjects from the field of business who were nominated by at least one of their selected classmates as a "most successful" graduate. Of these 50, 8 ranked in the upper one-fifth of their class academically, 9 in the second fifth, and 11 in each of the lower three quintiles. Kunkel concluded that there was no relationship between class academic rank and success in business.

Gambrill (1922), in his study of 1903 graduates of 11 colleges, found a nonsignificant .03 correlation between

academic achievement and salary (the criterion of career performance) for the 69 businessmen in her study.

Bridgman (1930) found that career performance defined as salary adjusted for number of years of work experience, could best be predicted by rank in college class. He studied 1,310 American Telephone and Telegraph Company male employees. Rank in college class proved to be a better predictor than the other two indices examined--campus achievement (i.e., leadership in editorial, forensic, managerial, social, athletic, musical, or dramatic activities) and early graduation.

Walters and Bray (1963) replicated Bridgman's earlier study at A.T.& T. using approximately 10,000 employees. The authors found rank in college graduating class to be a better predictor of salary adjusted for number of years of work experience, geographic region, and company department than either campus achievement or college quality.

Jepsen (1951) surveyed the 1929-1941 male graduates of Fresno State College. The respondents included 203 who were employed in business. Present (1948) salary and academic record yielded a nonsignificant -.05 correlation for the sample.

Williams and Harrell (1964) studied Stanford University business graduates (1927-1943). For 196 subjects, salary adjusted for length of time out of college

was cor
require
(MBA) p
MBA (.2
In addi
istrati
correla
MBA cou
None of

graduate
employed
undergra
of progr
in gener
leadersh
identifi
with the
.04; al

formance
as adapt
decisive
Two hund
liberal

was correlated with undergraduate GPA (.14), GPA for required courses in the Master of Business Administration (MBA) program (.13), and GPA for elective courses in the MBA (.22). The latter was the only significant correlation. In addition, 116 graduates completed self-reports of administrative level. This criterion of career performance was correlated with undergraduate GPA (.08), GPA for required MBA courses (.16), and GPA for elective MBA courses (.15). None of these three correlations was significant.

Pallett (1965) studied 184 University of Iowa graduates who were out of college from 5 to 10 years and employed in nontechnical jobs in business. He correlated undergraduate GPA with immediate supervisor's overall rating of progress and potential and eight components of "success" in general business--persuasiveness, drive, creativity, leadership, problem-solving ability, oral communication, identification with the business world, and identification with the company. The 10 correlations ranged from -.06 to +.04; all were nonsignificant.

Pallet and Hoyt (1968) correlated academic performance with the ratings of supervisors on such qualities as adaptability, creativity, problem-solving ability, decisiveness, dependability, communication, and leadership. Two hundred and sixty-six University of Iowa graduates in liberal arts and business administration from 1955 to 1959

showed no significant correlations, which ranged from -.34 to +.14. All the subjects had been out of college from five to ten years.

Selover (1970) studied 118 college graduates hired by the Prudential Insurance Company during 1956 and 1957 and found undergraduate GPA to be one of five measures which related well to rate of advancement within the company.

Cox (1971) used a sample of 453 students who during the 1962 and 1963 years attended trade, technical, and business schools. The subjects responded to a questionnaire on "vocational success"--i.e., job satisfaction, wage increases, job stability, job appropriateness. The elements of "vocational success" were correlated with instructor ratings, and all correlations were nonsignificant with the exception of a .30 correlation for a group of 131 business school females.

Harrell (1972) studied 434 Master of Business Administration (MBA) graduates from Stanford University five years after graduation. High earners in both large and small companies were found to have significantly higher GPAs than low earners.

Harrell and Harrell (1974) studied the 434 MBA-holders from Stanford University investigated in Harrell (1972) who had been out of Stanford for five years and an additional sample of 119 MBA-holders who had been out of college

in salar

lowest c

differen

the diff

second-y

optional

have inc

required

These re

companie

make mo

that th

increas

between

tion an

Georgia

tically

Master

pensati

years c

uates c

in jobs

college for ten years. They found significant differences in salary between the highest one-third in GPA and the lowest one-third in GPA. For five-year graduates the difference averaged $300, and for the ten-year graduates the difference averaged $700. The GPA predictor used was second-year MBA grades, which were typically obtained in optional courses in which it is assumed the student would have increased interest and motivation as compared with the required courses of the first-year in his MBA program. These results hold for subjects in both large and small companies. The authors conclude that MBAs with high grades make more money in business than MBAs with low grades and that the difference in earning power between the two groups increases as time passes.

Harrill (1973) found no significant correlations between undergraduate GPA and self-report of job satisfaction and job success for 77 male subjects who had attended Georgia Institute of Technology.

Weinstein and Srinivasan (1974) found a statistically significant correlation of .24 between graduate Master of Business Administration GPA and adjusted compensation (i.e., current salary, adjusted for number of years of work experience). The authors studied 116 graduates of Carnegie-Mellon University (1960-1966) who were in jobs such as accountant, supervisor, and manager.

<u>Engineering</u>

Rice (1913) studied graduates of Pratt Institute four to six years after they had completed their engineering degree. Correlations between grades and salary were computed for mechanical and electrical engineers within each of the three classes. Correlations ranged from .16 to .46 with two of the six found to be significant. Although methodologically sound, Rice's study suffers from salary being the sole criterion for career performance.

Twenty engineers were included in the Gambrill (1922) study of 11 colleges. Rank in class correlated with salary revealed no significant correlation.

Beatty and Cleeton (1928) correlated academic performance with two criteria of career performance--salary (.03) and rating of the importance of present position (.08). Both correlations were not significant. The subjects were 90 engineering graduates of Carnegie Institute of Technology (1923-1924) followed up in 1927.

Pierson (1947) had engineering faculty at the University of Utah rate 320 of the 1932-1941 graduates on occupational success. Engineering GPA correlated .43 with these ratings, and the author concluded that academic performance was a useful predictor of success in engineering. However, criterion contamination likely was a problem in that the engineering faculty rating the subjects on occupatio

academic

N

laborator

a study b

correlati

engineers

correlati

experienc

was perfo

the subje

These wei

salary.

T

engineers

on produc

career p

GPA for

correlat

but both

0.32 and

lations

more cre

creativi

degree.

occupational success were probably aware of the subjects' academic records.

Ninety-nine engineers employed in a research laboratory for Hughes Aircraft Company were subjects of a study by Martin and Pacheres (1962). A barely significant correlation between salary and college grades was found for engineers with four years of experience, and no significant correlation was obtained for those with six to eight years experience or for the total group. Furthermore, an analysis was performed in which weighted scores were developed from the subjects' college grades and reputation of college. These weighted scores did not correlate significantly with salary.

Taylor (1963) conducted a study in which 51 engineers and physicists were rated by their superiors on productivity and creativity. These two criteria of career performance were correlated with four-year GPA and GPA for the last two years of college. No significant correlation was found for either GPA with productivity, but both GPAs correlated significantly with creativity (.32 and .35, respectively). The GPA-creativity correlations may be spurious in that the subjects who were in more creative positions (and thus were rated higher in creativity) were those with education beyond the bachelor's degree. This postgraduate education experience could be

expected to correlate significantly with undergraduate GPA, thus possibly contaminating the results.

Taylor et al. (1963) report an anonymous study conducted at a large engineering research center in the late 1950's with 239 research engineers. A nonsignificant correlation of .06 was found between undergraduate GPA and merit rating of performance on research duties.

Muchinsky and Hoyt (1973) studied 127 graduates of Purdue University's College of Engineering (1956-1958 entering freshmen). The subjects had been out of college from 5 to 10 years. The authors correlated four measures of GPA (i.e., total college GPA, senior-year GPA, engineering core courses GPA, and the average of two senior-year design courses) with 15 measures of occupational success (i.e., supervisor's ratings on such qualities as scientific-technical knowledge, creativity-originality, practical judgment, and understanding of engineering problem-solving methodology). Of the 68 correlations examined, only six were statistically significant, and no correlation exceeded .30. The authors conclude that employers of engineering graduates should disregard GPA as a relevant variable in the hiring process.

Scientific Research

Science researchers have been the subject of a limited number of studies investigating the relationship

between
reporte
tories
perforn
class s

workinc
GPA was
tive of
3 of 1
writter
tional
of resc
formanc
to unde

physica
doctora
respond
career
(3) nur
society
best sc
naire f
on "sci

between academic and career performance. Mandell (1950) reported that 75 chemists in government research laboratories showed no relationship between ratings of on-the-job performance given by supervisors and colleagues and college class standing.

Taylor et al. (1961) studied 107 physical scientists working at two Air Force research centers. Undergraduate GPA was correlated with 14 criteria believed to be descriptive of the dimensions of "scientific contribution." Only 3 of 14 correlations were significant: productivity in written work (.27), creativity (.21), and current organizational status (.19). Quality of research work, originality of research work, scientific reputation, and overall performance were among the criteria which were not related to undergraduate GPA.

Harmon (1963) conducted a study in which 347 physical scientists and 157 biological scientists (all with doctoral degrees) employed by the Atomic Energy Commission responded to a questionnaire dealing with the following career information: (1) number of publications, (2) salary, (3) number and level of persons supervised, (4) scientific society memberships, (5) patents, and (6) self-ratings of best scientific or technical accomplishments. The questionnaire furnished the data by which judges rated each subject on "scientific competence." Undergraduate GPA correlated with "scientific competence" produced no significant correlations for six subgroups of the sample.

In another study Harmon (1964) had 355 National Science Foundation Fellowship candidates in five different fields of study rated by their colleagues as to the quality of their scientific accomplishments. When undergraduate GPAs were compared to these colleague ratings, no significant correlations (ranging from .00 to .20) were found with the exception of mathematicians, where the significant correlation of .49 was reported.

Ginzberg and Herme (1964) mailed a questionnaire to 548 fellowship winners who were in graduate or professional programs at Columbia University between 1944 and 1950. Career performance measures were income, academic rank, responsibility, quality of employing institution, and personal reputation. The only group for which prediction was reasonably accurate was the group composed of those students with the highest grades in their graduate courses. Students with the highest grades were more likely to be in the top career achievement category and less likely to be in the lowest career achievement category.

In a study by Taylor and Ellison (1964), 107 Air Force scientists were evaluated by peers and supervisors on 17 creativity measures--e.g., originality of written work, recent publications, overall performance, quality of research reports, and society memberships. No significant correlations were found for 13 of the 17 criteria when related to undergraduate GPA; slight positive correlations were found for the remaining four criteria. Possibly because they were primarily concerned with the study of creativity, Taylor and Ellison do not report the four statistically significant criteria.

Taylor and Ellison, in the same publication, reported the findings of their study in which 1,600 National Aeronautics and Space Administration scientists were rated by supervisors on creativity, overall performance, knowledge of work, initiative, judgment, industry, and number of publications and patents. Again, only slight positive correlations with undergraduate GPA were reported.

Chambers (1965) found a group of 213 psychologists and 225 chemists classified as creative to be significantly higher in undergraduate GPA than a matched control group of less creative subjects. However, methodological problems abound in Chambers' study. For instance, the two groups were not matched in terms of their interests. Also, undergraduate GPA data were gathered through self-report rather than through a records search. This is particularly worrisome since the median age of the entire sample was 53.

Finally, the studies by Martin and Pacheres (1962), Taylor (1963), and Taylor et al. (1963), cited in the

"Engineering" section, are relevant to this section. Either in whole or in part those three studies investigated engineers involved in scientific research activities.

Medicine

Before turning to studies involving physicians, two related research investigations involving nurses will be reported. Brandt and Metheny (1968) investigated 84 University of Washington School of Nursing graduates. Grades for clinical practice courses and grades for theory courses were correlated with self and supervisor ratings of professional competence. While grades in clinical practice courses were significantly correlated with both self and supervisor ratings, grades in theory courses were not significantly correlated with either self or supervisor ratings. The authors conclude that the low correlations which dominate the study support the position of little relationship between academic performance and professional performance for nurses.

Saffer and Saffer (1972) conducted a three-year study of 82 graduates of the Evanston (Illinois) Hospital School of Nursing. Self and employer ratings on 16 selected nursing activities (e.g., skill in general nursing procedures, ability to work efficiently and effectively under pressure, awareness of patient's physical needs and symptoms) were correlated with 12 measures of academic

perfor
the au
predic

career
no gen
mance.
resear
for ou
cation
subjec
subjec
recogn
perfor
appoin
cited
curren
clinic
to ind
perfor

future
by the
contai
quinti

performance. No significant correlations were found, and the authors concluded that academic performance in no way predicts on-the-job performance for nurses.

In regard to the relationship between academic and career performance among physicians, the literature reveals no general agreement as to the criteria for career performance. Criteria used to judge career performance vary from researcher to researcher and include: (1) peer nominations for outstanding performance, (2) salary, (3) board certification, (4) academic affiliation, (5) internist rating of subject's clinical competence, (6) number of journals subject subscribes to, (7) type of practice, (8) hospital recognition, (9) publications, (10) evaluation of internship performance by hospital official, and (11) professional appointments. As the above criteria indicate, the studies cited in this section often differ dramatically from the current study's criterion of career performance (i.e., clinical competence). Care has been taken in this section to indicate the particular researcher's criterion for career performance.

Kunkel's (1917) study cited earlier included 29 future physicians nominated to the "most successful" group by their classmates. Each of the first three quintiles contained about 25 percent of the sample, the fourth quintile contained 11 percent, and the fifth quintile contained 14 percent. Thus, there was little or no relationship between a "most successful" nomination and class rank.

Gambrill's (1922) study of 11 colleges included 30 physicians. Nonsignificant negative correlations were found when class rank was correlated with salary.

Peterson et al. (1956) had 88 North Carolina general practitioners rated by internists on six elements of general practice: clinical history, physical examination, use of laboratory aids, use of therapeutic measures, preventive medicine, and clinical records. The authors report that physicians in the upper 30 percent of their medical school graduating class received overall ratings of clinical competence (computed from the six elements of general practice) which were significantly higher than physicians in the middle 40 percent or the lower 30 percent. However, this relationship only held for young physicians (age 28 to 35) and not for older physicians. The results of this study support the suggestion of Wingard and Williamson (1973) that the relationship between medical school grades and clinical competence "wash out" as the experiences of the physician increase.

Richards et al. (1962) sampled 139 University of Utah College of Medicine graduates (1955 to 1958) and found the following significant correlations between medical school grades and evaluation of internship performance by hospital officials: (1) first-year GPA, .21; (2) second-year GPA, .24; and (3) third-year GPA, .45. These three correlations were adjusted for the quality of the hospital. The only unadjusted correlation which was significant was with third-year GPA (.35). No significant correlation (.03 and .06) was found between undergraduate GPA and evaluations of internship performance (adjusted or unadjusted).

Also at the University of Utah a series of studies was undertaken (Price et al., 1964; Taylor et al., 1964; Richards et al., 1965; and Taylor et al., 1965) which involved about 500 physicians. For each physician over 200 measures of performance were derived from interviews, directories and compendiums, faculty and alumni records, curriculum vita and bibliography, polled opinions of medical students, medical department chairmen, peers, questionnaires, and college transcripts. GPA was found to be unrelated to all measures of physician performance. Three GPA measures--undergraduate GPA, GPA for the first two years of medical school, and GPA for the last two years of medical school--were correlated with performance measures judged relevant for each of four subsamples--medical faculty (N = 102), board-qualified specialists (N = 190), urban general practitioners (N = 110), and rural-small town general practitioners (N = 105). Only 3 percent of these 849 subsample correlations were significant; 5 percent would be expected by chance.

Clute (1963) randomly chose 85 Canadian general practitioners and used internists' ratings of the subject's clinical competence as the criterion for career performance. His principal findings were: (1) in Ontario (N = 23) a significant correlation with academic performance of .56 and (2) in Nova Scotia (N = 39) a nonsignificant correlation with academic performance of .24.

Monk and Thomas (1970) used the criteria of type of practice, board certification, and academic affiliation with 724 of the 740 male graduates of the Johns Hopkins School of Medicine between 1948 and 1958. Physicians graduating in the top third of their class were twice as likely as bottom third graduates to pursue academic or research work. Those in the bottom third were more likely to be full-time practitioners. Among the full-time practitioners, the graduates with better grades were more likely to have board certification and academic affiliation. In summary, some of the correlations between career variables and class standing were significant, but none were very high--ranging from .17 to .29.

Westling-Wikstrand et al. (1970) used professional appointments and board certification as criteria with 81 female graduates of Johns Hopkins School of Medicine between 1948 and 1958. Among those with academic appointments two-thirds were ranked in the lower two-thirds of their class.

However, two-thirds of those not practicing medicine at all were ranked in the upper third of their class. Possibly the large number in the upper third of their class not practicing medicine represents a high percentage of medical researchers under the assumption that the researcher must function at a higher intellectual level than the practitioner.

Korman and Stubblefield (1971) used eight internship performance criteria in their study of 68 University of Texas Southwestern interns and consistently found no relationship between medical school grades and clinical performance.

Steiner et al. (1974) studied 142 physicians trained in the University of Toronto residency program during 1966 to 1971. Only 71 of the 142 (50 percent) qualified as psychiatric specialists by 1973, but all 15 in the upper third of their class qualified. In addition, 13 of 17 in the middle third qualified, but only 6 of 15 in the lower third of their class qualified. Data were available on only 47 subjects. The researchers' findings must be considered suspect in that the qualifying examination in psychiatry seems to be another academic performance measure rather than a measure of clinical competence or on-the-job performance. Thus, the qualifying examination results would be expected to correlate well with class standing.

Kegel-Flom (1975), investigating the 110 graduating males from the University of California School of Medicine at San Francisco (class of 1968), correlated medical school GPA with supervisor, self, and peer ratings of internship performance. Cumulative GPA was significantly correlated with supervisor ratings (.32), self ratings (.46), and peer ratings (.35).

Finally, at variance with Kegel-Flom is the work of Turner et al. (1974), who note that:

> Price and his associates (1971), having brought together a list of 87 positive qualities and 29 negative qualities that characterize the desirable and undesirable characteristics of physicians, discovered after two decades of research very few significant correlations between medical school grades and measures of career performance used in their study (p. 338).

Conclusion

In summary, a review of the literature dealing with the relationship between academic performance and career performance among professionals in teaching, business, engineering, scientific research, and medicine reveals little or no relationship between the two variables. Even where a statistically significant relationship between academic and career performance was found, methodological or criterion problems sometimes cast doubt on the validity of the findings. Many previous studies are not relevant due to criteria of career performance which are not meaningful in today's world. However, medical education has taken notable steps to better develop criterion measures of professional performance. The current study is, in fact, a part of this continuing effort.

An additional comment should be made in regard to statistical significance versus practical significance. Even in the studies where statistically significant correlations were reported, the practical significance of the reported correlations may be negligible. Issac and Michael (1971) urge researchers to ask whether statistically significant differences and correlations are educationally or practically significant. For example, in Richards et al.'s (1962) study cited earlier, the statistically significant correlation is .21 between first-year medical school grades and internship evaluations. However, Kerlinger and Pedhazur (1973) point out that the coefficient of determination (i.e., r^2 or the proportion of variance of one variable (e.g., career performance) "determined" by a second variable (e.g., academic performance) is equal to the square of the correlation coefficient (r). Thus, Richards et al.'s (1962) correlation coefficient of .21 is equal to an r^2 of .0441. Stated differently, less than 5 percent of the variance in career performance is "determined" by academic performance. Thus, nearly all the statistically significant correlations reported in this review of the literature can be brought into question as to their practical or educational significance.

Entry Level Predictors of Academic Performance and Clinical Competence

When investigating the question of entry levels as predictors of academic performance and clinical competence, we are essentially asking whether aptitude predicts achievement. Block (1971), in a summary of the research results dealing with the general aptitude-achievement question concludes that

> aptitudes, as measured by either standardized aptitude tests or simple pretests of a student's prior knowledge of a subject, are predictive of . . . the level to which a pupil will learn in a given time [p. 92].

Bloom (1968) recognizes the relationship between aptitude and achievement but believes that with a student population randomly distributed on aptitude, 95 percent of the students will achieve mastery or competence level if each learner receives optimal quality of instruction and the learning time he requires.

Turning to medical education, although recent studies (Elstein et al., 1973; McGuire and Page, 1973) suggest that clinical competence is a domain-specific capability, investigations have not been undertaken to examine what aptitudes relate to clinical competence in the various domains or body systems (e.g., neuromuscular, cardiovascular, genital-urinary, gastrointestinal) which compose the typical medical education program. Rather, researchers in the past have approached the question of entry levels as predictors of academic performance or clinical competence primarily from an admissions standpoint. In an effort to improve selection/admission procedures, administrators and researchers have focused on the factors which might best predict the criterion of medical school success (which operationally usually means grades).

According to *Medical School Admission Requirements, U.S.A. and Canada 1975-76*, the principal admissions criteria of medical schools in the United States and Canada are premedical GPA, scores on the Medical College Admission Test (MCAT) with its four subtests (Verbal Ability, Quantitative Ability, General Information, Science), psychological/ personality factors, personal interviews, and letters of recommendation. The literature contains an abundance of studies relevant to the first two criteria (premedical GPA and the MCAT) as entry level predictors of medical school success, a moderate amount on psychological/personality factors, but a dearth of research studies on the personal interview and the letter of recommendation. Thus, a survey of the literature on premedical GPA, the MCAT, and psychological/personality factors as entry level predicto

criterio

was medi

clinical

current

Studies
of Medic
Clinica

Carolina

lation

between

Each ph

trained

rating

clinica

aids, t

records

Tulane

found t

yielded

randomly

found n

supervis

predictors will follow. It should be noted that the criterion of interest in most of the studies undertaken was medical school success (i.e., grades), rather than clinical competence, a variable of more interest in the current study.

Studies Using the MCAT as a Predictor of Medical School Grades and/or Clinical Competence

Peterson et al. (1956) had data from only 30 North Carolina general practitioners but found no positive correlation between MCAT scores and medical school grades or between MCAT scores and physician clinical performance. Each physician was rated on his clinical performance by trained observers who spent three to three and one-half days rating the physicians on their skill in the following areas: clinical history, physical examination, use of laboratory aids, therapy, preventive medicine, and keeping of clinical records.

Hoffman et al. (1963), using 1,278 students from Tulane University School of Medicine over a 12-year period, found that MCAT scores and clinical grades in medical school yielded consistently low correlations.

Howell and Vincent (1967), in their study of 54 randomly selected U.S. Public Health Service physicians, found negative correlations between MCAT scores and supervisors' ratings of clinical performance.

Bartlett (1967) followed the progress of 49 University of Rochester medical students (1949-1962) through medical school and the early phases of their professional careers. Bartlett found that low MCAT scorers, especially in the Verbal Ability and Science subtests, did not differ from medical students with higher MCAT scores in either medical school academic performance or career performance.

For 235 University of Kentucky medical students (1965-1968), Haley (1973) found that MCAT scores and medical school GPA did not significantly correlate. However, when grades were factor analyzed, MCAT scores correlated significantly with two factors: the Behavioral Science Factor and the Genetics-Physiology Factor. No significant correlations were found between MCAT scores and the factors entitled Anatomy Factor, Medical Factor, or General Science Factor.

Studies Using the MCAT and Premedical GPA as Predictors of Medical School Grades and/or Clinical Competence

Holt and Luborsky (1958) conducted a massive study in which 247 psychiatric residents who trained at the Menninger School of Psychiatry between 1946 and 1951 were investigated. Both MCAT scores and premedical GPA failed to predict supervisors' ratings of intern performance for this large group of psychiatric residents.

Hill (1959) studied 1,000 medical students from the years 1950-1957 at the State University of New York College of Medicine in Brooklyn and found significant positive relationships between premedical GPA and academic performance in the first year of medical school and between the MCAT and academic performance in the first year of medical school.

Johnson (1962), using 927 applicants for admission to the State University of New York, Upstate Medical Center, for classes entering 1956-1960, found premedical GPA, adjusted by the mean MCAT at the student's undergraduate college, to be the best single predictor of medical school academic success.

Buehler and Trainer (1962) used male students admitted to the University of Oregon Medical School from 1949 to 1954 who could be categorized into the top and bottom 10 percent of their class on the MCAT, premedical GPA, and medical school grades. After completing a variety of intercorrelations, the researchers concluded that premedical GPA and the MCAT are good predictors of medical school academic performance.

At St. Louis University School of Medicine, Roemer (1965) studied 863 students from the graduating classes of 1955 through 1963. He found that premedical GPA correlates with medical school grades (with a range of .272 for fourth year mec

first tw

of the M

with mec

in that

optimal

the hig

ently,

subtest

of the

He thus

satisfa

medical

medical

MCAT an

was for

endorse

an undi

and MCA

school

school.

Alabama

of the

year medical school grades to .427 for the average of the first two years of medical school) better than the subtests of the MCAT or the average of the MCAT subtests correlate with medical school grades. However, he was disappointed in that even with the multiple regression equation which optimally utilizes the MCAT subtests and premedical GPA, the highest multiple R obtained was .534. Stated differently, the best combination of the predictors (the MCAT subtests and premedical GPA) only accounts for 28.5 percent of the variance in the criterion (medical school grades). He thus concludes that premedical GPA and the MCAT are not satisfactory as predictors for the purpose of selection into medical school. The best prediction was for first year medical school grades (correlated with the subtests of the MCAT and premedical GPA at .534) and the worst prediction was for fourth year medical school grades. Finally, Roemer endorses premedical GPA as a better predictor than the MCAT.

Fredericks (1967) studied 82 freshman students at an undisclosed mid-western school and found premedical GPA and MCAT scores not to be significantly related to medical school academic performance during the first year of medical school.

Stefanu and Farmer (1971) studied 166 University of Alabama medical students and found that the Science subtest of the MCAT is a good predictor of first-year medical school success (.43 correlation between MCAT-Sci and first-year medical school grades) for students with high premedical GPAs. However, no correlation between the MCAT-Sci and first-year medical school grades was found for students with low premedical GPAs.

Moffatt et al. (1971), in a study using four entering classes at the University of Iowa College of Medicine, found no positive correlations between MCAT scores and academic performance in gross anatomy. However, premedical GPA was a good predictor of gross anatomy performance.

Best et al. (1971), using 399 University of Illinois students, found that performance on multiple-choice examinations during medical training is best predicted by an approximately equal weighting of premedical GPA, type of college attended, quantitative MCAT, and science MCAT. However, these four predictors or any other predictors used in the admissions procedure were not reliable in forecasting student performance on Patient Management Problems (PMP) or instructor rating of clinical clerkship performance. Thus, the authors conclude that premedical GPA and parts of the MCAT are better predictors of a measure of medical school grades (i.e., the multiple choice examination) than of measures of clinical competence (i.e., PMP performance or instructor clinical clerkship rating).

Kegel-Flom (1975), cited earlier, found that for 110 male interns none of the four subtests of the MCAT was significantly correlated with supervisor, self, or peer ratings of internship performance (correlations ranged from -.21 to +.22). Furthermore, in regard to premedical GPA, neither cumulative GPA, science GPA, nor GPA for the last two semesters correlated significantly with supervisor, self, or peer ratings (correlations ranged from .03 to .16).

Discussion of the MCAT and Premedical GPA as Predictors

The studies cited above seem to lead to an equivocal conclusion as to the effectiveness of the MCAT and premed-GPA as predictors of medical school success. After reviewing representative studies in the literature, Gough et al. (1963) concluded that the MCAT has low predictive validity when the criterion is medical school grades or completion of training and no predictive validity when the criterion is performance in internship or professional practice. Furthermore, the authors conclude that premedical GPA is a promising predictor but more so for the basic sciences than the clinical sciences or internship. Miller (1961) maintains that the most consistent single predictor of medical school academic achievement is premedical GPA with correlations ranging from .27 to .60. Johnson and Hutchins (1966), in their massive study of attrition in medical school, endorse premedical GPA and the MCAT as valuable predictors of medical school success or failure. Erdmann et al. (1971), in their review of the MCAT, notes the strong relationship between MCAT scores and medical school grades, but in another publication, Erdmann (1972) observes that although the MCAT has been a factor in reducing the medical school attrition rate from 11 percent in 1962 to 5 percent in 1972, it has not been successful in predicting performance in clinical settings.

Two measurement limitations interfere with the ability of the MCAT or premedical GPA to predict the criterion of either medical school grades or clinical competence. First, studies involving admissions data, including those cited in this review of the medical literature, invariably suffer from the phenomenon of restriction of range (Miller, 1971; Gough et al., 1963; Magnusson, 1967). The range of scores on the predictor (i.e., either premedical GPA or the MCAT) are restricted in that subjects were not admitted to medical school unless they reached a certain cutoff point; those below the cutoff point were excluded from medical school. Thus, the entire range of scores on the predictor variable was not available. This restriction of range results in a lowered correlation coefficient between the predictor and the criterion (either medical school grades or clinical competence), that is, less predictive ability for the MCAT or premedical GPA. On the other hand, if all medical school applicants who completed the MCAT and had a premedical GPA were allowed to attend medical school (thus eliminating the restriction of range), it would be expected that the MCAT and premedical GPA would be better predictors of both medical school grades and clinical competence. How much better, of course, is speculative, and no conclusions in regard to the studies cited can be drawn.

Second, Miller (1961) notes that reported low correlations between MCAT scores and medical school grades may, in part, be due to the unreliability of the criterion, medical school grades. In investigating the correlation between the MCAT and medical school grades, we are examining criterion-related validity, a coefficient whose size is dependent upon the reliability of both the predictor and the criterion. The criterion-related validity coefficient will decrease in size to the extent that either the predictor or the criterion is unreliable. Since the subtests of the MCAT consistently have reliability coefficients above .90, the predictor can be minimally faulted in regard to unreliability. However, the unreliability of medical school grades is well known. Although Miller does not mention either premedical GPA as a predictor or clinical competence as a criterion, both of which are subject to unreliable measurement, correlations involving either of these two variables would be lowered by the unreliability of their measurement.

Personality Factors

Holt and Luborsky (1958), in a study cited earlier in which 247 psychiatric residents who trained at the Menninger School of Psychiatry between 1946 and 1951 were investigated, conducted two analyses involving the personality variables of the subjects. In the first instance, two expert judges rated 64 of the psychiatric residents on 32 personality variables after having reviewed the assessment file of each subject. An assessment file for a subject contained his tests, interview evaluations, credentials, etc. Those psychiatric residents rated by their supervisors as high in overall competence were characterized by personality variables which could be grouped under the headings of (1) effective in human relationships, (2) substantial psychological health, (3) kindly and understanding toward others, and (4) adequate emotional control.

In the companion analysis, Holt and Luborsky extracted from the 247 psychiatric residents the upper 13 percent and lower 13 percent in overall competence and had the 33 subjects in each group rated by their supervising psychiatrists on 40 personality variables. In summary, the residents in the upper group were found to be more intelligent, more sensitive, more independent in thinking and judgment, and warmer but more self-contained and even-tempered. In addition, the upper group residents expressed themselves more appropriately, had better relationships with their patients, co-workers, and supervisors, and were more interested in their work and in learning. Finally, the better residents were more stable, more mature, and had greater insight into themselves.

Morris (1958), who used the Thurstone Temperament Schedule, was an early advocate of personality tests as a means of predicting academic performance in medical school. Working at the University of Iowa, Morris promoted the study of personality factors which are indicative of successful and unsuccessful medical students in the hope of increasing predictive ability in regard to medical school performance.

Kole and Matarazzo (1965), using the Edwards Personal Preference Schedule with 80 University of Oregon medical students, found the most dominant needs of the typical medical student to be the need to persevere with difficult tasks and the need to work to the best of one's ability. Conversely, medical students were relatively low in the need for emotional support.

Johnson and Hutchins (1966) report the findings from a study of 2,812 medical students (class of 1960).

In comparison with medical students who made normal academic progress, academic dropouts were less theoretical and aesthetic (thus less creative) and less competitive but more practical, materialistic, and religious. Academic dropouts were more conforming, neat, and compulsive but less achievement-oriented and less aggressive than normal progress students. The authors used the Allport-Vernon-Lindzey Study of Values and the Edwards Personal Preference Schedule.

Rothman and Flowers (1970), who studied 174 University of Toronto freshman medical students, using the Personality Research Form and the Sixteen Personality Factor Questionnaire, found first-year students with higher grades to have a greater need for achievement, endurance, social recognition, and understanding, and to be, in general, aloof, intelligent, serious, cautious, and intellectually curious and diverse.

Rothman (1973), in a study of the complete four years of undergraduate medical training for the sample used in the Rothman and Flowers study (N = 160 with 14 subjects lost to attrition), found that during the first year of medical school training high achievers could be differentiated from low achievers by greater needs for achievement, endurance, and introversion, three traditionally learning-related needs. However, during the last two years, high achievers were differentiated from low achievers by status-associated traits--need for dominance, need for social recognition, and exhibition.

Haley et al. (1971), in a study combining both cognitive and noncognitive factors among 991 medical students from diversified medical schools, found that high MCAT scorers differed from low MCAT scorers in being more independent and more aesthetic but less conforming, less religious, and less dogmatic. In addition to the MCAT, the authors used the Allport-Vernon-Lindzey Study of Values, the Survey of Interpersonal Values, and Rokeach's Form E Dogmatism Scale.

With a sample of 114 University of Kentucky students, Haley and Lerner (1972) found that, in general, medical students who achieve the best grades in basic science courses prior to clinical experiences tend to be submissive, uncritical of authority, cynical, ambitious to obtain personal, political, or economic power, persevering, less intelligent, less socially concerned, and more flexible in their approach to life in comparison with classmates whose grades are not as high. The authors utilized the Allport-Vernon-Lindzey Study of Values and the Edwards Personal Preference Schedule.

Turner et al. (1974) examined the relationship between clinical competence (as derived from ratings of videotaped workups of the student's communication skill, interpersonal skill, and physical examination skill) and 60 personality variables (as derived from the Opinion, Attitude, and Interest Survey, the Omnibus Personality Inventory, the Myers-Briggs Type Indicator, and the Sixteen Personality Factor Questionnaire). The subjects were 50 third-year Ohio State University medical students. The researchers found that students with good clinical skills were those who (1) felt sufficiently adequate and stable to admit discomfort under stress, (2) were open and spontaneous in response to first-hand experience, and (3) were high on impulse expression.

In a study of medical school success, Gough and Hall (1975) studied 1,071 medical students who entered the University of California School of Medicine in San Francisco between 1955 and 1968. Of the 1,071 subjects, 1.6 percent became academic dropouts and 3.8 percent became nonacademic dropouts. Using 50 cognitive and noncognitive variables derived from the MCAT, premedical GPA, ratings of admissions interviewers, the Adjective Check List, and the California Psychological Inventory, the authors developed multiple regression equations to predict medical school completion or withdrawal. They also stress the importance of both cognitive and noncognitive factors in medical school prediction efforts. In regard to personality factors the nondropouts as compared to the dropouts were found to be more independent and less conforming.

Finally, Kegel-Flom (1975), in a study cited earlier in this review, administered the California Personality Inventory to 110 subjects (male interns). Two significant findings were reported: (1) peers who rated an intern high in performance saw him as significantly more tolerant than an intern they rated low in performance; and (2) interns who rated themselves high in performance saw themselves as more dominant and inflexible than interns who rated themselves low in performance.

Summary of Personality Characteristics of the Successful Medical Student/Physician

The studies reviewed indicate in general that the medical student/physician who achieves high medical school grades is persevering, creative, aesthetic, competitive, achievement-oriented, aggressive, intelligent, intellectually curious and diverse, and independent. He also tends to be less conforming, less materialistic, and less religious than his classmate who does not achieve as well.

In addition, the medical student/physician who is rated highly as a clinician is, in general, tolerant, sensitive, congenial, mature, stable, spontaneous, nondefensive, self-insightful, intelligent, independent, self-contained, and interested in his work and in learning.

Summary of the Chapter

This chapter dealt with three principal topics: (1) the concept, origin, and development of competency-based instruction, (2) the relationship of academic performance to career performance, and (3) entry levels as predictors of academic performance and clinical competence. Competency-based instruction was defined with emphasis on two prime components--performance objectives and criterion-referenced tests. The origins of competency-based instruction in teacher education was noted, and specific developments at four medical colleges were detailed. Studies were cited to evidence that typical measures of academic performance frequently have little or no relationship to career performance among professionals in the fields of teaching, business, engineering, science research, and medicine. Important studies and authorities interested in entry levels based on the Medical College Admission Test and premedical GPA were cited. It is generally conceded that while premedical GPA and the MCAT are good predictors of academic performance in medical school, they have not been successful in predicting clinical performance. Research investigating the personality characteristics of successful and unsuccessful medical students was cited, and the personality characteristics of the successful medical student/physician were summarized.

CHAPTER III

DESIGN OF THE STUDY

Sample

In this study the sample of subjects is identical to the available population. Eighty-seven students enrolled in Michigan State University's College of Osteopathic Medicine's (MSU-COM) Neuromuscular Instructional System for first-year students during the summer term, 1975, were used. In actuality, 83 subjects were strictly first-year students who were in the process of completing their first calendar year at MSU-COM. Four subjects were out of sequence in their Doctor of Osteopathy (D.O.) program and, although enrolled for the first-year Neuromuscular Instructional System, had been a MSU-COM student for more than one calendar year. Nevertheless, for the purpose of this study, all 87 students were considered first-year osteopathic medical students in that they were enrolled for the Neuromuscular Instructional System for the first time.[1]

[1]Two students enrolled for the instructional system were not included in the study. One of the two students dropped the instructional system early in the term while the second student was repeating the instructional system and thus did not meet the criterion of being a student in the Neuromuscular Instructional System for the first time.

Sample Characteristics

Kempthorne (1961) referred to the available population as the "experimentally accessible population" and to the populations which contained similar characteristics and to which the researcher might want to generalize as "target populations." The "experimentally accessible population" is the MSU-COM entering class of 1974. The following description of the entering class of 1974 will be useful to the reader or researcher wishing to generalize to his "target populations" which are similar in regard to demographic characteristics. Cornfield and Tukey (1956) have advanced the practice of generalizing from one's sample to populations composed on similar characteristics. Thus, the results of the current study might be generalized to analogous populations of medical students.

The demographic data below include the 84 students for which the MSU-COM Office of Admissions had accumulated information.

Sex:	57 males and 27 females	
Age:	Range = 20 to 44; Median = 23; Mean = 24.82	
Race:	White = 66	Black = 13
	Spanish-surnamed = 3	Oriental = 2
Program:	All students were enrolled in the regular Osteopathic Medical Program.	

Residency:

Michigan	= 67	Indiana	= 1
Alabama	= 2	Massachusetts	= 1
New Jersey	= 2	Maine	= 1
Ohio	= 2	New York	= 1
Arizona	= 1	Texas	= 1
California	= 1	Washington	= 1
Connecticut	= 1	Wisconsin	= 1
Illinois	= 1		

Degrees earned:

Associate	= 14	Baccalaureate	= 78
Master's	= 15	Doctorate	= 4

Majors in which degrees were earned--36 different areas with the following occurring more than once:

biology	= 35	liberal arts	= 3
zoology	= 10	microbiology	= 3
psychology	= 8	agronomy	= 2
premedicine	= 5	anthropology	= 2
biological sciences	= 4	business	= 2
pharmacy	= 4	engineering	= 2
physiology	= 4	mathematics	= 2
chemistry	= 3	nursing	= 2

Schools at which degrees were earned--48 different schools with the following occurring more than once:

Michigan State University	= 25
Wayne State University	= 14
University of Michigan	= 12
University of Detroit	= 4
Lansing Community College	= 3
University of Wisconsin	= 3
Aquinas College	= 2
Central Michigan University	= 2
Kalamazoo Valley Community College	= 2
Michigan Technological University	= 2
Oakland Community College	= 2
Oakland University	= 2
Taylor University	= 2
Western Michigan University	= 2

Medical College Admission Test Scores (N = 81):

	Mean	Median
Verbal	503.30	515
Quantiative	535.00	545
General Information	515.25	515
Science	524.75	535

Premedical Grade-Point Average (GPA):

	Mean	Median
Science	3.07	3.02
Non-Science	3.13	3.17
Overall	3.09	3.07

Threats to Generalization

Campbell and Stanley (1963) have listed four threats to generalization or, to use their term, external validity. First, the "reactive effect of testing" occurs when a pretest affects subjects' responsiveness to the experimental variable and thus makes results unrepresentative of populations not pretested. This threat can be discounted for the current study in that a single experimental variable was not involved. The study was concerned with a variety of student performance scores diffused over a period of time and not practicably influenced by the one measure which could be classified accurately as a pretest, the test of cognitive knowledge in neurology, a one-half hour, 35-item, multiple-choice test. Data for the other entry level variables--the measures of history-taking, physical examination, and doctor-patient relations skill--were gathered in the context of course instruction and student evaluation for a course grade.

Second, the "interaction of selection biases and the experimental variable" occurs when the sample was not randomly selected. Thus, the subjects in the sample used may respond differently from the population from which they were selected preventing generalization to the "experimentally accessible population" and consequently to the "target populations." This threat is not relevant to this current study in that no selection procedure, random or otherwise, was involved. All available subjects ($N = 87$) were included in the study.

Third, the "reactive effect of experimental arrangements" occurs when the experimental setting affects results such that similar results would not have occurred in a non-experimental setting. This threat can be discounted in that the only deviation from the normal course of instruction for the Neuromuscular Instructional System was the aforementioned pretest. In essence, there was no experimental setting. Data was gathered in the context of the normal instruction and evaluation.

Fourth, "multiple-treatment interference" occurs when subjects receive more than one treatment and thus are affected by previous treatments when undergoing a new experimen

relevant

experimen

E

Campbell

to gener

categori

Populati

Kempthor

accessib

concerne

experime

The exp

study h

its ade

target

validit

herein

treatm

ecolog

curren

1.
2
3.
4
5

experimental treatment. Again, this threat is not relevant in that the current study contained no traditional experimental variable.

Bracht and Glass (1968) have extended the work of Campbell and Stanley (1963) and classify various threats to generalization (i.e., external validity) under two categories--"population validity" and "ecological validity." Population validity has two subcategories, one dealing with Kempthorne's (1961) distinction between an "experimentally accessible population" and a "target population" and another concerned with an interaction effect between different experimental treatments and different levels of subjects. The experimentally accessible population for the current study has been described, and the reader must decide upon its adequacy in regard to generalization to any specific target population. Bracht and Glass's second population validity subcategory is not relevant because the research herein reported deals neither with traditional experimental treatments nor different levels of subjects.

Bracht and Glass detail ten subcategories of ecological validity, all of which pose no threat to the current study:

1. Hawthorne effect
2. novelty and disruption effects
3. multiple-treatment interference
4. pretest sensitization
5. failure to explicitly describe the independent variable

6.

7.
8.
9.

10.

A numb

parall

effect

the re

in Cam

Multip

Pretes

of tes

is rel

variab

within

relevan

is unwa

group o

have be

there i

and tre

repeate

6. failure to explicitly describe the dependent variable
7. posttest sensitization
8. history effects
9. interaction of time of measurement and treatment effects
10. experimenter's expectations and behaviors.

A number of Bracht and Glass's ecological threats have parallels in Campbell and Stanley (1963). The Hawthorne effect and novelty and disruption effects are similar to the reactive effect of experimental arrangements mentioned in Campbell and Stanley's account of external validity. Multiple-treatment interference is mentioned in both texts. Pretest sensitization is paralleled by the reactive effect of testing in the work of Campbell and Stanley.

Bracht and Glass point out that ecological validity is reliant upon explicit descriptions of the independent variable and the dependent variable. This will be done within the present chapter. Posttest sensitization is not relevant in that no posttests were included in the study.

Bracht and Glass's concern about history effects is unwarranted in light of the fact that there was only one group of subjects in the study and not two groups who could have been differently affected by extraneous events. Also, there is no basis for an interaction of time of measurement and treatment effects, for the current study was not a repeated measures design.

Finally, Bracht and Glass recognize the threat to ecological validity of the experimenter's expectations and behaviors. In this regard, while the subjects were aware that the researcher was collecting data for a study, they were unaware of the purpose and content of the study. In addition, the researcher's function as an administrative aide allowed him to be viewed by the subjects principally as an instructional assistant rather than a researcher. Therefore, the researcher believes the subjects were not influenced by any possible experimental expectations. The researcher's contact with subjects was within an administrative context, and he was not involved in active instruction of subject matter nor in the scoring of any measure or test. Thus, it is reasonable to assume that the researcher's behavior did not influence the data.

In summary, using the criteria of Campbell and Stanley (1963) and Bracht and Glass (1968), the researcher believes no threats to external validity (i.e., generalization) exist. Of course, the reader must decide for his individual circumstances as to the justification for generalizing from the current experimentally accessible population to his particular target population.

Description of the Neuromuscular Instructional System

The Neuromuscular Instructional System for first-year osteopathic medical students was the context within which the data for the current study were gathered. It therefore is important to describe the Neuromuscular Instructional System as experienced by MSU-COM's students. Much of the following has been prepared by Lawrence Jacobson, D.O., Coordinator of the Neuromuscular Instructional System. (Adaptations have been made for the sake of consistency with the present text.)

Introduction to the Neuromuscular Instructional System

The Neuromuscular Instructional System provides an integrated approach to the teaching of the normal and abnormal functions of the nervous system. The system incorporates both basic science and clinical science principles that are pertinent to clinical problems involving the nervous system. The primary effort of the system is directed at providing the student with a basic functional comprehension of those principles that are pertinent to the diagnosis, evaluation, and treatment of neurological and musculoskeletal disorders.

Objectives

The objectives of the system are divided into two basic areas: (1) didactic concepts and principles and (2) diagnostic and clinical skills.

1. Didactic concepts and principles
    a. To develop comprehension of those basic science concepts and principles that are pertinent to the understanding of the diagnosis and treatment of clinical disorders of the nervous system.
    b. To develop an understanding of specific clinical disorders involving the nervous and musculoskeletal systems with major emphasis on specific conditions encountered with great frequency in medical practice.
2. Diagnostic and clinical skills
    a. To develop those motor skills necessary for conducting a competent and organized neurological examination.
    b. To develop ability in differentiating normal from abnormal neurological findings and to localize the site(s) of neuropathology on the basis of these abnormal clinical signs.
    c. To develop comprehension of the procedural technique, the rationale, and the use of

specialized tests pertinent to the diagnosis of the following neurological problems:

1) Electromyography
2) Nerve conduction time
3) Chronaxie times
4) Muscle biopsy
5) Myelogram
6) Cerebral angiogram
7) Pneumoencephalogram and ventriculogram
8) Electroencephalogram
9) Spinal tap and cerebrospinal fluid analysis
10) Echoencephalogram and ophthalmodynamemetry
11) Brain scan and EMI computor tomography
12) Skull and spinal x-rays
13) Drug provocative tests (e.g., Tensilon test, Histamine provocative test)
14) Cold caloric testing.

Instruction

1. Classroom teaching consisted of four hours per day, five days per week for the ten weeks of the term. Classroom teaching integrated basic science material with practical clinical input utilizing lectures, demonstrations, focal problem conferences, motion pictures, television tapes, and other audiovisual aids.

2. Self-instructional laboratories were conducted for the first five weeks of the term, one laboratory experience per week. A student was responsible each week for viewing the assigned television tape, practicing the fundamental motor skills for a neurological assessment of patients which the tape presented, and completing a self-assessment examination dealing with the content of the tape.

a mea

pract

Durin

compl

inati

made

oppor

Gradi

studen

obtain

T-scor

measur

relate

weighte

the thr

muscula

one tel

test we

satisfa

instruc

and foca

3. Real and simulated patients were utilized as a means of providing the student with the opportunity to practice and improve his neurological clinical skills. During weeks six through nine each student completed two complete neurological evaluation history and physical examinations with simulated patients. In addition, each student made one trip to a hospital or clinic where he had the opportunity to observe and examine neurological patients.

Grading

MSU-COM has a Pass/No Grade system of evaluating students. For the Neuromuscular Instructional System to obtain a Pass grade, a student had to achieve (1) a weighted T-score of 35 for the four multiple-choice tests which measured basic and clinical science concepts and principles related to the Neuromuscular Instructional System (each test weighted 25 percent) and (2) a weighted T-score of 35 for the three tests of clinical skills related to the Neuromuscular Instructional System (two TV case evaluations and one televised PMP) and the neuroanatomy practical test (each test weighted 25 percent). In addition, the student had to satisfactorily complete all assignments for the self-instructional laboratories, patient contact experiences, and focal problems.

Measures Used in the Study

This section is concerned with the measures used in the study. Before the measures are described in detail, they will first be introduced. The variable grades was determined by weighting the eight measures used in the Neuromuscular Instructional System as follows:

Grading Devices	Percent of Overall Grade
Neuroanatomy mid-term	16.67
Neuroanatomy final	16.67
Neurosciences mid-term	16.67
Neurosciences final	16.67
Practical Neuroanatomy	8.33
TV Case Evaluation I	8.33
TV Case Evaluation II	8.33
Patient Management Problem	8.33
	100.00

Clinical competence was composed of two variables: (1) ability to perform a complete neurological evaluation history and physical examination (see Appendix B for examples) and (2) patient rating of student in performing a complete neurological evaluation history and physical examination (see Appendix C). Below is the point distribution for the parts of the measure used to assess the ability to perform a complete neurological evaluation history and physical examination (NEHPE):

Section	Points for Section
Chief complaints	3
Onset and course of chief complaints	6
Past history--family, medical, social	5
Systems review	2
Physical examination	17.25
Summary	4.25
Provisional diagnosis	1
Differential diagnosis	6
Tests	4
Therapy	2
	50.50

Below are the semantic differential scales used to measure the patient rating of the student in performing a complete neurological evaluation history and physical examination (PR-NEHPE):

seven semantic differential scales with 5 rating points

1	2	3	4	5
very interested				very uninterested
did not use uncommon words or concepts				frequently used uncommon words or concepts
clear				unclear
very understanding				not understanding
secure				insecure
patient				impatient
very gentle during the physical examination				very rough during the physical examination

Four entry levels were assessed: cognitive knowledge in neurology (see Appendix D), history-taking skill (HIST-PD), physical examination skill (PE-PD), and doctor-patient relationship skill (DPR-PD). The latter three entry levels were assessed using the measure found in Appendix F. Below are listed the four entry levels and the manner in which these data were collected:

Measure of Entry Levels	Data Collected
cognitive knowledge in neurology (COG) see Appendix D	During the first week the student took the 35-item multiple-choice test.
history-taking skill (HIST-PD) see Appendix F	Prior to the outset of the Neuromuscular Instructional System, the student was rated by a member of the medical faculty as the student performed a complete history and physical examination on a simulated patient.
physical examination skill (PE-PD) see Appendix F	Same as history-taking skill
doctor-patient relationship skill (DPR-PD) see Appendix F	Same as history-taking skill

Measures Used to Obtain Grades

Of the eight grading devices which composed the variable grades, the first five which follow were measures of concepts and principles related to the Neuromuscular Instructional System. The last two grading devices which follow were measures of clinical skills related to the Neuromuscular Instructional System.

1. Neuroanatomy Mid-Term Test.--This test was weighted as one-sixth (16.67%) of the overall weighted T-score. The test was composed of 63 multiple-choice neuroanatomy items.

2. Neuroanatomy Final Test.--This test was weighted as one-sixth (16.67%) of the overall weighted T-score. The test was composed of 65 multiple-choice neuroanatomy items.

3. Neurosciences Mid-Term Test.--This test was weighted as one-sixth (16.67%) of the overall weighted T-score. The test was composed of 102 multiple-choice items from the following areas: physiology, biochemistry, pathology, pediatrics, microbiology, pharmo-therapeutics, neurology, and community medicine.

4. Neurosciences Final Test.--This test was weighted as one-sixth (16.67%) of the overall weighted T-score. The test was composed of 105 multiple-choice items from the following areas: physiology, biochemistry, pathology, pediatrics, microbiology, pharmo-therapeutics, neurology, and community medicine.

5. Practical Neuroanatomy Test.--This test was weighted as one-twelfth (8.33%) of the overall weighted T-score. The test was a one-hour, 50-item test composed of two parts: (1) a part containing magnified slide projections

of brain sections in which the student must perform such tasks as (a) identifying the brain structure with its blood supply and the fibers which lead to the structure and (b) indicating what cardinal signs and clinical deficits would be produced by a particular brain lesion; and (2) a part containing real brains (whole and sections) in which the student must perform such tasks as (a) identifying certain blood vessels, (b) indicating the possible clinical problem which could result from the occlusion of a certain blood vessel, (c) identifying the clinically important structures of the brain, and (d) indicating the cardinal functions of the brain structures.

6. Internal consistency reliability coefficients.--Kuder-Richardson Formulas 20 and 21 were used to estimate the internal consistency reliability for the measures of concepts and principles related to the Neuromuscular Instructional System. Table 1 shows the reliability of these five measures.

7. TV Case Evaluations.--Each of these two tests were weighted as one-twelfth (8.33%) of the overall weighted T-score. The student viewed a physician conducting a neurological evaluation history and physical examination and then had to complete an answer sheet asking for appropriate diagnosis, laboratory tests and procedures, and management (see Appendix A).

Table 1. Internal consistency reliability for measures of concepts and principles related to the Neuromuscular Instructional System

Test	Reliability Coefficient	Formula
Neuroanatomy Mid-Term	.9068	Kuder-Richardson 20
Neuroanatomy Final	.8794	Kuder-Richardson 20
Neurosciences Mid-Term	.7840	Kuder-Richardson 20
Neurosciences Final	.8555	Kuder-Richardson 20
Neuroanatomy Practical	.8324	Kuder-Richardson 21[a]

[a]The Neuroanatomy Practical was hand-scored. Thus, the percentage of students correctly answering each item could not practicably be tabulated. Consequently, the Kuder-Richardson 21 Formula, which requires only the mean of the test scores, the number of items for the test, and the test vraiance, was used. Mehrens and Lehmann (1973) and Ebel (1972) both note that the Kuder-Richardson 21 Formula underestimates the reliability coefficient when the test items vary in difficulty. Since it can be assumed that the Neuroanatomy Practical items did vary in difficulty, the reported KR-21 value is an underestimate of the reliability coefficient.

8. Televised Patient Management Problem (PMP).--This test was weighted one-twelfth (8.33%) of the overall weighted T-score. The student viewed a physician conducting a neurological history and physical examination and then was asked to make a diagnosis, indicate appropriate laboratory techniques and procedures, and recommend proper treatment and management using the "erasure test" (PMP) format in which the examinee makes sequential decisions based on immediate feedback received from the answer sheet.

Measures Used to Obtain Clinical Competence

1. Measure of the ability to perform a complete neurological evaluation history and physical examination (NEHPE).--This was a measure completed by each student after he had finished a complete neurological evaluation history and physical examination with a simulated patient. The measure (see Appendix B) has ten parts. The measure totals to 50.5 points, each of the ten parts having been appropriately weighted using the rational judgment of the researcher and two physician's assistants.[2] Below is a brief description of the criteria used in assessing the student's responses for the ten parts with the points assigned to each part.

---

[2]The use of physician's assistants in the Neuromuscular Instructional System and the current study will be described in Chapter V.

a. Chief Complaints (3 points). Assessment was based on the clarity and accuracy with which the student detailed the two or three basic neurological problems described by the patient.

b. Onset and Course of Chief Complaints (6 points). Assessment was based on the clarity, accuracy, and organization of the chief complaints in regard to their onset, location, duration, severity, course of previous treatment, and other general symptoms and descriptive characteristics.

c. Past History--Family, Medical, Social (5 points). Assessment was based on the accuracy of the student's responses in regard to family history (1 point), medical history of patient (3½ points), and social history (½ point). See Appendix B for details of Past History.

d. Systems Review (2 points). Assessment was based on the accuracy with which the student identified relevant problems in the patient's neuromuscular system or other body system in addition to the chief complaints.

e. Physical Examination (17¼ points). Assessment was based on the accuracy with which the student reported the findings for his patient. The sections of the physical examination were general appearance

(½ point), general findings (1½ points), mental status (¼ point), speech and language function (¼ point), reflexes (5 points), sensory (2½ points), muscle function and gait (2 3/4 points), cerebellar and dorsal column function (1¼ points), extrapyramidal (1¼ points), and cranial nerves (2 points).

f. Summary (4¼ points). Assessment was based on the accuracy with which the student identified the general neurological results (¼ points), the areas of neurological system dysfunction (1½ points), and the anatomical location of neurological system dysfunction (2½ points).

g. Provisional or Working Diagnosis (1 point). Assessment was based on the accuracy of the student's responses in identifying the principal categorical diagnosis and the associated specific disorder. In this section the student must apply his problem-solving skills, clinical judgment, and decision-making skills to the data he has gathered in the neurological history and physical examination.

h. Differential Diagnosis (6 points). Assessment was based on the accuracy of the student's responses in identifying the specific types of pathology which he hypothesizes as possible causes of the patient's problems. As in the provisional diagnosis, the assi

(a)

phys

disa

two

not

exam

as e

ical

entl

the

scor

105

student must apply his problem-solving skills, clinical judgment, and decision-making skills to the data he has gathered.

i. Tests (4 points). Assessment was based on the accuracy with which the student identified laboratory tests and other diagnostic procedures appropriate to his differential diagnosis.

j. Therapy (2 points). Assessment was based on the appropriateness of the specific and supportive therapy recommended by the student.

In scoring each examination, the physician's assistant (PA) used a key which he had formulated from (a) his own complete neurological evaluation history and physical examination of the simulated patient and (b) the disabilities with which the simulated patient was programmed.

A measure of inter-rater reliability between the two raters (PAs) was not computed since the two PAs were not familiar with the other's keys (i.e., they had not examined the simulated patients of the other PA). However, as evidence that the two were scoring the complete neurological evaluation history and physical examinations consistently, it should be noted that the mean difference between the two PAs was 1.35 on a 50.50 scale. That is, PA 1 scored 73 exams with a mean score of 39.71 and PA 2 scored 105 exams with a mean score of 38.36. A test of equality of means w

of ranc

circums

mean d

nonran

icance

scorin

physic

for th

the me

neuro

by th

(a) t

accep

Apper

neuro

in th

(d)

of t

This

Tinn

and

Each

means was not performed since the necessary assumption of random assignment was not possible in the instructional circumstances. However, in the researcher's judgment, a mean difference of 1.35 points between the two PAs for two nonrandom groups has no practical or educational significance. Furthermore, it should be noted that prior to the scoring of the examinations, the researcher and the two physician's assistants carefully determined the criteria for the assigning of points for each of the ten parts of the measure.

The measure of the ability to perform a complete neurological history and physical examination is considered by the researcher to be a valid and reliable measure in that (a) the ten-part form used to obtain the measure is a well-accepted procedure, (b) the neurological cases selected (see Appendix B) for programming were typical and important neurological cases, (c) the twenty simulated patients used in the study were carefully trained and standardized, and (d) a mean difference of 1.35 points between the ratings of the two PAs was insignificant in a practical sense. This measure was first used for research purposes by Tinning (1973).

For the complete neurological evaluation history and physical examinations, simulated patients were utilized. Each student examined two different simulated patients.

Tin

use

use

add

the

obt

who

Sol

med

who

simu

Appe

comp

exam

neur

the

sema

deve

Hess

simu

pati

rate

words

Tinning's (1973) experimental study demonstrated that the use of simulated patients is a viable alternative to the use of real patients in assessing clinical competence. In addition, support for the use of simulated patients within the context of measuring clinical performance can be obtained from the work of (1) Barrows and Abrahamson (1964), who assessed skills in clinical neurology, (2) McGuire and Solomon (1971), who assessed clinical skills in orthopedic medicine, and (3) Jason et al. (1970) and Froelich (1969), who dealt with doctor-patient interaction skills.

The three specific neurological cases for which simulated patients were programmed can be found in Appendix B.

2. Patient rating of student in performing a complete neurological evaluation history and physical examination (PR-NEHPE).--After the completion of a complete neurological evaluation history and physical examination, the simulated patient rated the student using a 7-item semantic differential evaluative scale (see Appendix C) developed from Tinning (1973), Turner et al. (1972), and Hess (1969). Each student was rated by two different simulated patients. This form provided a measure of doctor-patient relationship skill; the items asked the patient to rate the student's interest in the patient, use of uncommon words, clarity of explanation, sympathy-empathy, personal

secur

neuro

as an

the p

compl

exami

Measu

(COG)

measu

befor

The t

The i

was .

purpo

physi

relat

Instr

(PD)

durin

histo

The s

security, patience, and gentleness in conducting a neurological physical examination.

A coefficient alpha reliability value was computed as an estimate of the measure's internal consistency. Using the patient ratings of students in performing their second complete neurological evaluation history and physical examination, the coefficient alpha value was .8203.

Measures Used to Obtain Entry Levels

1. Test of cognitive knowledge in neurology (COG).--This was a 35-item test (see Appendix D) used to measure the student's cognitive knowledge in neurology before exposure to the Neuromuscular Instructional System. The test was administered during the first week of the term. The internal consistency reliability (Kuder-Richardson 20) was .6178. This measure was first used for research purposes by Tinning (1973).

2. Ratings of history-taking skill (HIST-PD), physical examination skill (PE-PD), and doctor-patient relationship skill (DPR-PD).--Prior to the Neuromuscular Instructional System and as part of a physical diagnosis (PD) course taken by first-year osteopathic medical students during winter term, 1975, the students performed a complete history and physical examination on a simulated patient. The simulated patients all had real disabilities which were

reco

by t

pati

rate

to l

taki

four

The

Marc

(197

hist

recorded along with other important physical problems by the COM faculty physicians who examined the simulated patients before the student exams. A supervising physician rated the student on a 5 (superior: needs no supervision) to 1 (inadequate or omitted) scale for the five history-taking elements, seven physical examination elements, and four doctor-patient relationship elements (see Appendix F). The ratings were gathered during late February and early March 1975.

The following statement is from Tinning and Yinger (1975) in regard to the data obtained from the student's history and physical examinations:

> Each student's Practical Examination Rating Sheet was scored to yield a raw score total. Early in the process it was obvious that there was considerable difference among instructors in the criteria that they were using to determine a good performance. The mean raw percentage score that was given by all raters ranged from a high of 95.5% to a low of 63.75%. In order to correct for this difference and to standardize the scores as much as possible, a constant was computed based on the difference between an instructor's mean for each item and the mean for all instructors for the same item. This constant was then either added to a student's score (if the instructor was below the mean) or subtracted from a student's score (if the instructor was above the mean).
>
> After standardization, the mean total score . . . was 78.89%, ranging from a high of 97.65% to a low of 67.58%. . . . The reliability (internal consistency) of the scale was computed on the standardized scores yielding a coefficient alpha of .6835 [p. 21].

sc

Calendar for Data Collection

The data were gathered according to the following schedule:

- Prior to the Neuromuscular Instructional System

  late February to early March 1975
  entry level ratings of history-taking skill, physical examination skill, and doctor-patient relationship skill.

- During the Neuromuscular Instructional System

  June 18-24, 1975:
  entry level test of cognitive knowledge in neurology

  July 7:
  TV Case Evaluation I

  July 9-17:
  entry level neurological psychomotor practical test

  July 17:
  Neuroanatomy Mid-Term Test

  July 21:
  Neurosciences Mid-Term Test

  July 23 to August 5:
  first complete neurological evaluation history and physical examination and first patient rating of student performance

  July 25:
  TV Case Evaluation II

  August 5:
  Practical Neuroanatomy Test

  August 6-19:
  second complete neurological evaluation history and physical examination and second patient rating of student performance

1

2

3

hypoth

August 12:
Neuroanatomy Final Test

August 15:
Televised Patient Management Problem

August 25:
Neurosciences Final Test.

Null Hypotheses

Three null hypotheses were investigated.

1. *For first-year osteopathic medical students in a neuromuscular instructional system, clinical competence in performing a complete neurological evaluation history and physical examination is not related to grades.*

2. *For first-year osteopathic medical students in a neuromuscular instructional system, clinical competence in performing a complete neurological evaluation history and physical examination is not related to entry levels.*

3. *For first-year osteopathic medical students in a neuromuscular instructional system, grades are not related to entry levels.*

Analysis of the Data

The data analytic procedures for the null hypotheses were as follows:

compe

peter

two c

exami

stude

grade

used

Systen

compet

compet

penden

of cog

taking

(d) rat

Null Hypothesis	Data Analytic Procedure	Dependent Variable(s)	Independent Variable(s)
1	multivariate regression	NEHPE PR-NEHPE	Grades
2	multivariate multiple regression	NEHPE PR-NEHPE	COG HIST-PD PE-PD DPR-PD
3	multiple regression	Grades	COG HIST-PD PE-PD DPR-PD

For Hypothesis 1 (no relationship between clinical competence and grades) the dependent variable clinical competence has two elements: (a) mean T-score on performing two complete neurological evaluation history and physical examinations and (b) mean T-score of the patient ratings of student performance on (a). The independent variable was grades--an overall weighted T-score for the eight measures used for grading purposes in the Neuromuscular Instructional System.

For Hypothesis 2 (no relationship between clinical competence and entry levels) the dependent variable clinical competence has the same two elements cited above. The independent variable entry levels has four elements: (a) test of cognitive knowledge in neurology, (b) rating of history-taking skill, (c) rating of physical examination skill, and (d) rating of doctor-patient relationship skill.

a

c

as

Mu

the

For Hypothesis 3 (no relationship between grades and entry levels) the dependent variable is grades as cited above and the independent variable is entry levels as cited above.

Scheifley and Schmidt's (1973) adaptation of Finn's Multivariance was the statistical package used to analyze the data for the three hypotheses.

sepa

data

inde

resu

inte

are

tist

Null

Data

the n

CHAPTER IV

RESULTS OF THE ANALYSIS

In this chapter the three hypotheses are discussed separately. For each hypothesis, the null hypothesis, the data analytic procedure, the dependent variable, and the independent variable are stated. For each hypothesis the results of the analysis are reported, and finally the interpretations following from the results of the analysis are discussed. The analysis uses .05 as the level of statistical significance for all results in each hypothesis.

Hypothesis One

The relationship between clinical competence and grades.

Null Hypothesis

For first-year osteopathic medical students in a neuromuscular instructional system, clinical competence in performing a complete neurological evaluation history and physical examination is not related to grades.

Data Analytic Procedure

Multivariate regression analysis was used to test the null hypothesis.

In

mea

Res

sig

gra

tio

NER

lat

gra

var

sig

wit

sin

sta

bee

Dependent Variable--Clinical Competence

1. Mean T-score for two complete neurological evaluation history and physical examinations (NEHPE).
2. Mean T-score for two patient ratings of two complete neurological evaluation history and physical examinations (PR-NEHPE).

Independent Variable--Grades

Overall weighted T-score derived from the eight measures used for grading (see pp. 80-84).

Results

Table 2 shows that there is a statistically significant relationship between clinical competence and grades ($p < .0001$). Thus, the null hypothesis of no relationship between clinical competence and grades is rejected. NEHPE is correlated with grades .4565 and PR-NEHPE is correlated with grades .2582. While the independent variable grades would be significantly related to either dependent variable (NEHPE is significant at $p < .0001$ and PR-NEHPE is significant at $p < .0158$), a multivariate regression equation with PR-NEHPE as a dependent variable is not appropriate since the relationship between PR-NEHPE and grades is statistically insignificant ($p < .0762$) after NEHPE has been controlled. This is also apparent from the F values for

down

has

cont

grad

par

as

Hyp

an

dep

Int

bet

in

sig

com

the

are

ana

and

mus

con

As

for the step down Fs. NEHPE entered first has a step down F of 22.3803, but PR-NEHPE, when added after NEHPE, has a step down F of only 3.2233.

Thus, the complete regression equation would contain NEHPE and PR-NEHPE as dependent variables and grades as the independent variable. However, the most parsimonious regression equation would eliminate PR-NEHPE as a dependent variable. Table 3 reveals the results of Hypothesis 1 using this regression equation. Table 4 is an intercorrelation matrix for all independent and dependent variables of Hypothesis 1.

Interpretations

In the first study to examine the relationship between clinical competence and grades among physicians in the classroom-laboratory portion of their training, a significant correlation of .4565 has been found. For the competency-based instructional model in medical education the implication of this significant correlation is in the area of student evaluation devices. The results of the analysis support the utilization of both the grading devices and the clinical competence measures used in the Neuromuscular Instructional System in the development of a competency-based instructional model in medical education. As this study has shown that the student's ability to

perfor
physic
the st
evalua
signif
clinic
course
device

compe
corre
studi
That
resea
that
corre
compe
becau
inte
atti

perform a complete neurological evaluation history and physical examination (NEHPE) and the patient ratings of the student's ability to perform a complete neurological evaluation history and physical examination (PR-NEHPE) were significantly related to grades, it is likely that similar clinical competence measures in other medical education courses would be significantly related to the grading devices used in those courses.

The significant correlation between clinical competence and grades supports the argument that the low correlation between these two variables reported in other studies is due to the "intervening experiences hypothesis." That is, Wingard and Williamson (1973) speculate and the research of Peterson et al. (1956) supports the position that medical school classroom-laboratory grades do not correlate well with career performance (with clinical competence as the most important component of this variable) because of the effects of intervening experiences such as internships, residencies, and office practice on habits, attitudes, and interpersonal skills.

Table 2. Test of the null hypothesis: no relationship between clinical competence and grades

F	Degrees of Freedom	p Less Than
13.0945	2 and 84	.0001*

Dependent Variable	R^2	R	F	p Less Than	Step Down F	p Less Than
NEHPE	.2084	.4565	22.3803	.0001*	22.3803	.0001*
PR-NEHPE	.0667	.2582	6.0699	.0158*	3.2233	.0762

*Significant.

Table 3.

R^2
.2084

Table 4.

NEHPE
PR-NEHP
Grades

Table 3. Test of the null hypothesis: no relationship between clinical competence (measured only by NEHPE) and grades

R^2	R	F	p Less Than	Degrees of Freedom
.2084	.4565	22.3803	.0001*	1 and 85

*Significant.

Table 4. Intercorrelation matrix for all independent and dependent variables of Hypothesis 1

	NEHPE	PR-NEHPE	Grades
NEHPE	1.0000		
PR-NEHPE	.1983	1.0000	
Grades	.4565	.2582	1.0000

Null Hyp

Data Ana

to test

Depende

1.

2.

Indepen

1.

2.

3.

Hypothesis Two

The relationship between clinical competence and entry levels.

Null Hypothesis

For first-year osteopathic medical students in a neuromuscular instructional system, clinical competence in performing a complete neurological evaluation history and physical examination is not related to entry levels.

Data Analytic Procedure

Multivariate multiple regression analysis was used to test the null hypothesis.

Dependent Variable--Clinical Competence

1. Mean T-score for two complete neurological evaluation history and physical examinations (NEHPE).
2. Mean T-score for two patient ratings of two complete neurological evaluation history and physical examinations (PR-NEHPE).

Independent Variable--Entry Levels

1. Test of cognitive knowledge in neurology (COG).
2. Rating of history-taking skill from a physical diagnosis course (HIST-PD).
3. Rating of physical examination skill from a physical diagnosis course (PE-PD).

4.

Result

signif

entry

relati

is rej

and th

p<.00

the fc

p<.69

two de

step d

into t

dent v

test i

in the

reveal

(signi

remair

as pre

4. Rating of doctor-patient relationship skill from a physical diagnosis course (DPR-PD).

Results

Table 5 shows that there is a statistically significant relationship between clinical competence and entry levels ($p < .0270$). Thus, the null hypothesis of no relationship between clinical competence and entry levels is rejected. The multiple correlation (R) between NEHPE and the four independent variables is .3946 (significant, $p < .0072$); the multiple correlation between PR-NEHPE and the four independent variables is .1623 (insignificant, $p < .6962$).

Table 6 contains the univariate F tests for the two dependent variables (i.e., NEHPE and PR-NEHPE) and the step down F tests for PR-NEHPE after NEHPE has been entered into the regression equation. With PR-NEHPE as the dependent variable of interest, no univariate F or step down F test is significant. Thus, only NEHPE need be included in the most parsimonious regression equation.

Turning to the independent variables, Table 6 reveals that only HIST-PD (significant, $p < .0096$) and PE-PD (significant, $p < .0361$) are significantly related to the remaining dependent variable (i.e., NEHPE).

Evidence for the importance of HIST-PD and PE-PD as predictors can be seen in Table 7. HIST-PD, when entered second into the regression equation, accounts for an additional 7.5026 percent of the variation in NEHPE. PE-PD, when entered third after COG and HIST-PD into the regression equation, accounts for an additional 4.6479 percent of the variation in NEHPE.

Thus, while the complete regression equation is significant ($p < .0270$) with NEHPE and PR-NEHPE as dependent variables and COG, HIST-PD, PE-PD, and DPR-PD as independent variables, the most parsimonious regression equation would include only one dependent variable (i.e., NEHPE) and two independent variables (i.e., HIST-PD and PE-PD). Table 8 reveals the results of Hypothesis 2 using this regression equation. Table 9 is an intercorrelation matrix for all independent and dependent variables of Hypothesis 2.

Interpretation

A significant correlation of .3946 has been found between clinical competence and entry levels. For the competency-based instructional model in medical education the implications of this finding are in the areas of screening, placement, and individualized instruction. History-taking skill (HIST-PD) and physical examination skill (PE-PD) were found to be significantly correlated with the student's ability to perform a complete neurological evaluation history and physical examination (NEHPE). The correlations were .2580 for HIST-PD and .1942 for PE-PD.

Thus, these preliminary findings suggest that at least two entry levels and possibly others can be used in a neurology course for screening, placement, and individualized instruction. Entry level prerequisites could be established for screening whether or not a student is permitted to enroll for a neurology course. In a similar manner, entry level measures could be used to place qualified students within a neurology course. For instance, a student who excels in history-taking skill (e.g., a physician's assistant who has had abundant clinical experience) might spend less time with history-taking and more time in an area where he needs study or practice. Finally, entry level measures could serve a valuable individualized instruction function by directing students toward the areas of neurology in which they need assistance. For example, one student may do poorly on the neuroanatomy portion of a cognitive entry level measure while another student may be "all thumbs" in performing neurological physical examinations. Entry level measures could be used as early cues to these students in alerting them of their deficiencies.

In a like manner, evaluative research might reveal that similar entry level measures can be used for screening, placement, and individualized instruction in other medical education courses. Various cognitive, psychomotor, and affective entry level measures might be developed for use

Table 5. Test of the null hypothesis: no relationship between clinical competence and entry levels

F	Degrees of Freedom	p Less Than
2.2411	8 and 162	.0270*

Dependent Variable	R^2	R	F	p Less Than	Step Down F	p Less Than
NEHPE	.1557	.3946	3.7802	.0072*	3.7802	.0072*
PR-NEHPE	.0263	.1623	.5546	.6962	.8410	.5032

*Significant.

Table 6. Univariate F tests and step down F tests for Hypothesis 2

		NEHPE		PR-NEHPE		Step Down F for PR-NEHPE After NEHPE Has Been Entered First	
Step[a]	Independent Variable	Univariate F	p Less Than	Univeraite F	p Less Than	Univariate F	p Less Than
1	COG	2.5776	.1121	1.2267	.2712	.6599	.4189
2	HIST-PD	7.0373	.0096*	.5885	.4452	.7037	.1955
3	PE-PD	4.5435	.0361*	.0128	.9104	.3408	.5610
4	DPR-PD	.4615	.4989	.4302	.5138	.6833	.4109

[a]Each step is the F-test of the indicated independent variable with the preceding independent variables controlled.

*Significant.

Table 7. Percent of additional variance accounted for in NEHPE by including the independent variables in the regression equation in a stepwise manner

Step	Independent Variable	Percent of Additional Variance Accounted for in NEHPE
1	COG	2.9432
2	HIST-PD	7.5026
3	PE-PD	4.6479
4	DPR-PD	.4752

Table 8. Test of the null hypothesis: no relationship between clinical competence (measured only by NEHPE) and entry levels (measured only by HIST-PD and PE-PD)

R^2	R	F	p Less Than	Degrees of Freedom
.1037	.3220	4.8585	.0101*	2 and 84

*Significant.

Table 9. Intercorrelation matrix for all independent and dependent variables of Hypothesis 2

	NEHPE	PR-NEHPE	COG	HIST-PD	PE-PD	DPR-PD
NEHPE	1.0000					
PR-NEHPE	.1983	1.0000				
COG	.1716	.1193	1.0000			
HIST-PD	.2580	-.0928	-.0867	1.0000		
PE-PD	.1942	-.0251	-.1110	.0058	1.0000	
DPR-PD	.0186	.0444	-.0073	.2740	.0500	1.0000

in courses dealing with such body systems as the cardiovascular, the respiratory, the gastrointestinal, and the urinary systems.

Hypothesis Three

The relationship between grades and entry levels.

Null Hypothesis

For first-year osteopathic medical students in a neuromuscular instructional system, grades are not related to entry levels.

Data Analytic Procedure

Multiple regression analysis was used to test the null hypothesis.

Dependent Variable--Grades

Overall weighted T-score derived from the eight measures used for grading (see pages 80-84).

Independent Variable--Entry Levels

1. Test of cognitive knowledge in neurology (COG).
2. Rating of history-taking skill from a physical diagnosis course (HIST-PD).
3. Rating of physical examination skill from a physical diagnosis course (PE-PD).
4. Rating of doctor-patient relationship skill from a physical diagnosis course (DPR-PD).

<u>Results</u>

Table 10 shows that there is a statistically significant relationship between grades and entry levels ($p < .0044$). Thus, the null hypothesis of no relationship between grades and entry levels is rejected. The multiple correlation (R) between grades and the four independent variables is .4086.

Turning to the independent variables, Table 11 reveals that only COG ($p < .0004$) is significantly related to the dependent variable grades. Similarly, Table 12 shows that COG, when entered into the regression equation first, accounts for 14.2791 percent of the variation in the dependent variable. The remaining three independent variables add little to the total variance when entered after COG.

Thus, while the complete regression equation is significant ($p < .0044$) with grades as the dependent variable and COG, HIST-PD, PE-PD, and DPR-PD as independent variables, the most parsimonious regression equation would include only COG as an independent variable. Table 13 reveals the results of Hypothesis 3 using this regression equation. Table 14 is an intercorrelation matrix for all independent and dependent variables of Hypothesis 3.

Interpretation

A significant correlation of .4086 has been found between grades and entry levels. This null hypothesis was not tested for its immediate relevance to the competency-based instructional model. Rather the null hypothesis of no relationship between grades and entry levels was tested (1) to complete the testing of all meaningful regression equations utilizing the three variables of interest--clinical competence, grades, and entry levels--and (2) to furnish data to medical educators who utilize more traditional grading devices and who want to introduce entry level measures for screening, placement, and individualized instruction purposes. In the "Interpretation" section to Hypothesis 2 (see pages 104-105), the use of entry level measures to (1) screen admission into a medical education course, (2) differentially place students within a course, and (3) cue individualized instruction was described when clinical competence was the criterion measure. In an analogous manner, entry levels could be used with grades as the criterion measure.

The analysis for Hypothesis 3 reveals the student's cognitive knowledge in neurology (COG) to be the only retained entry levels of the four tested. That is, for all practical purposes, grades can be predicted as well with just COG as with all four entry levels. This is not a surprising result in that both COG and the dependent variable (grades) are heavily academic in nature.

Table 10. Test of the null hypothesis: no relationship between grades and entry levels

R^2	R	F	p Less Than	Degrees of Freedom
.1669	.4086	4.1083	.0044*	4 and 82

*Significant.

Table 11. F-tests for the independent variables of Hypothesis 3 when entered into the regression equation in a stepwise manner

Step	Independent Variable	F-Test for the Independent Variable with the Preceding Independent Variables Controlled	p Less Than
1	COG	14.1590	.0004*
2	HIST-PD	1.5039	.2235
3	PE-PD	.9036	.3446
4	DPR-PD	.0009	.9757

*Significant.

Table 12. Percent of additional variance accounted for in grades by including the independent variables in the regression equation in a stepwise manner

Step	Independent Variable	Percent of Additional Variance Accounted for in Grades
1	COG	14.2791
2	HIST-PD	1.5077
3	PE-PD	.9070
4	DPR-PD	.0010

Table 13. Test of the null hypothesis: no relationship between grades and entry levels (measured only by COG)

R^2	R	F	p Less Than	Degrees of Freedom
.1428	.3779	14.1590	.0004*	1 and 85

*Significant.

Table 14. Intercorrelation matrix for all independent and dependent variables of Hypothesis 3

	Grades	COG	HIST-PD	PE-PD	DPR-PD
Grades	1.0000				
COG	.3779	1.0000			
HIST-PD	.0896	-.0867	1.0000		
PE-PD	.0522	-.1110	.0058	1.0000	
DPR-PD	.0328	-.0073	.2740	.0500	1.0000

Summary of the Three Hypotheses

Table 15 is the summary table for the three null hypotheses. All three null hypotheses are rejected, and it is concluded that there are statistically significant relationships between (1) clinical competence and grades, (2) clinical competence and entry levels, and (3) grades and entry levels. The three regression equations for testing the null hypotheses would be useful for prediction of (1) clinical competence from grades, (2) clinical competence from entry levels, and (3) grades from entry levels. Furthermore, more parsimonious regression equations were developed. For Hypothesis 1, NEHPE as the dependent variable and grades as the independent variable, for Hypothesis 2, NEHPE as the dependent variable and HIST-PD and PE-PD as the independent variables, and for Hypothesis 3, grades as the dependent variable and COG as the independent variable were found to be the most parsimonious regression equations.

Additional Analysis of the Data

In addition to the three hypotheses of interest, further analysis of the data was undertaken. Three topics will be discussed: (1) using grades* as a variable by "weighting into" the grading devices a neurological psychomotor practical test (PSYMO); (2) examining the predictive ability of the test of cognitive knowledge in neurology

Table 15. Summary table

Null Hypothesis	p Less Than	Dependent Variable(s)	Independent Variable(s)	Dependent Variables Retained After Step Down Fs Performed	Independent Variables Retained After Stepwise Regression Performed	p Less Than
No relationship between clinical competence and grades	.0001*	NEHPE PR-NEHPE	Grades	NEHPE	Grades	.0001*
No relationship between clinical competence and entry levels	.0270*	NEHPE PR-NEHPE	COG HIST-PD PE-PD DPR-PD	NEHPE	HIST-PD PE-PD	.0101*
No relationship between grades and entry levels	.0044*	Grades	COG HIST-PD PE-PD DPR-PD	Grades	COG	.0004*

*Significant.

(COG) in regard to individual grading devices and the two subsets of grading devices; and (3) examining the predictive ability of the two subsets of grading devices in regard to clinical competence.

The Neurological Psychomotor Practical Test Weighted into the Grading Devices

The neurological psychomotor practical test (PSYMO) was a practical test of the student's ability to perform the psychomotor tasks required for a neurological physical examination (see Appendix E). For each task the student was rated as excellent (3), adequate (2), or failing (1) in the three areas of (a) accuracy of instructions, (b) positioning of patient, and (c) technique. The total point score was the student's neurological psychomotor practical test score. The test was administered by two physician's assistants, who underwent careful training to assure that they were using the same criteria for rating the student. The test was administered during the fourth and fifth weeks of the term and prior to the student's first complete neurological evaluation history and physical examination with a simulated patient. This measure was first used for research purposes by Tinning (1973).

Since PSYMO could have been used as a grading device in the Neuromuscular Instruction System, it was decided to "weight in" PSYMO and examine the two null hypotheses involving grades utilizing the new variable (grades*). Dr. Lawrence Jacobson, the instructional system coordinator, judged that if PSYMO were to be used as a grading device, it should be assigned a weight equal to the four multiple-choice tests. Thus, the weights of the individual grading devices for the new variable grades* were:

Grading Device	Percent of Total Grades
Neuroanatomy Mid-Term	14.28
Neuroanatomy Final	14.28
Neurosciences Mid-Term	14.28
Neurosciences Final	14.28
PSYMO	14.28
TV Case Evaluation I	7.15
TV Case Evaluation II	7.15
Patient Management Problem	7.15
Practical Neuroanatomy	7.15
	100.00

Table 16 shows a comparison of the multiple correlations (R) and the levels of significance utilizing both grades and grades* for Hypothesis 1 (no relationship between clinical competence and grades) and Hypothesis 3 (no relationship between grades and entry levels). It can be concluded from Table 16 that utilizing grades* rather than grades would be of no advantage. The decisions in regard to statistical significance are not affected. Inspection of the multiple correlations (R) reveals that only a slight gain in predictive ability results when grades* is substituted for grades in Hypothesis 1 and that a loss in predictive ability results when grades* is substituted for

Table 16. Hypotheses 1 and 3 tested with Grades*

	Independent Variable(s)	Dependent Variable(s)	R	Level of Significance for Each Dependent Variable	Level of Significance for the Hypothesis
Hypothesis 1: No relationship between clinical competence and grades	Grades	NEHPE	.4565	.0001*	
		PR-NEHPE	.2582	.0158*	.0001*
	Grades*	NEHPE	.4587	.0001*	
		PR-NEHPE	.2990	.0050*	.0001*
Hypothesis 3: No relationship between grades and entry levels	COG HIST-PD PE-PD DPR-PD	Grades	.4086	.0044*	.0044*
	COG HIST-PD PE-PD DPR-PD	Grades*	.3664	.0177*	.0177*

*Significant.

grades in Hypothesis 3. In addition, it should be noted that for Hypotheses 1 and 3 the same variables drop out of the most parsimonious regression equation when grades* is entered as with grades (see Summary Table, page 116).

The Test of Cognitive Knowledge in Neurology as a Predictor of Grading Devices

Hypothesis 3 revealed the test of cognitive knowledge in neurology (COG) to be the only entry level which significantly predicted grades. Thus, it was decided to examine how well COG predicted (1) each individual grading device, (2) those five grading devices which were designed to measure the concepts and principles related to the Neuromuscular Instructional System, and (3) those three grading devices which were designed to measure clinical skills related to the Neuromuscular Instructional System. Below the eight grading devices are listed with the weights each received when it was included in either the concepts and principles subset or the clinical skills subset:

Grading Device	Weight as a Concepts and Principles Measure	Weight as a Clinical Skills Measure
Neuroanatomy Mid-Term	22.222	--
Neuroanatomy Final	22.222	--
Neurosciences Mid-Term	22.222	--
Neurosciences Final	22.222	--
Practical Neuroanatomy	11.112	--
TV Case Evaluation I	--	33.333
TV Case Evaluation II	--	33.333
Patient Management Problem	--	33.334
	100.000	100.000

Table 17 shows that COG correlated significantly with the four major measures of concepts and principles in the Neuromuscular Instructional System (i.e., Neuroanatomy Mid-Term and Final and Neurosciences Mid-Term and Final) but did not correlate with the minor measure of concepts and principles (i.e., Practical Neuroanatomy). A slightly significant correlation between COG and TV Case Evaluation I was found while the other two measures of clinical skills in the Neuromuscular Instructional System (i.e., TV Case Evaluation II and the PMP) were not significantly correlated with COG. These results are not unexpected. The test of cognitive knowledge in neurology is basically a concepts and principles measure and is constructed in the multiple choice format. On the other hand, the clinical skills measures claim to involve application, an area somewhat different from concepts and principles.

However, it is most interesting that when the three individual measures of clinical skills are combined to form one variable, COG correlates significantly with this composite variable ($r = .2382$; $p < .0263$). It appears that the three measures form a three-item test of clinical skills and, by doing so, become a more reliable measure of clinical skills in the same way in which the reliability of a test (and hence its predictive value) is increased with the addition of new test items.

Tab

Gra
Neu
Neu
Neu
Neu
Pra
TV
TV
Pat
Con
Cli

ski

com

nif

fou

spe

Gra
Cli

rel

exa

con

mea

Table 17. Correlation between grading devices and the test of cognitive knowledge in neurology (COG)

Grading Device	Correlation (r) with COG	p Less Than
Neuroanatomy Mid-Term	.2511	.0190*
Neuroanatomy Final	.2862	.0072*
Neurosciences Mid-Term	.3424	.0012*
Neurosciences Final	.3113	.0034*
Practical Neuroanatomy	.0812	.4545
TV Case Evaluation I	.2147	.0459*
TV Case Evaluation II	.0985	.3639
Patient Management Problem	.1142	.2922
Concepts and Principles Measures	.3264	.0021*
Clinical Skills Measures	.2382	.0263*

*Significant.

Furthermore, this finding that individual clinical skills measures do not correlate well with COG but that a composite of the three clinical skills measures does significantly correlate is consistent with PMP research which found physician clinical skills to be case- or domain-specific (Elstein et al., 1973; McGuire and Page, 1973).

Grading Devices as Predictors of Clinical Competence

Hypothesis 1 revealed that grades were significantly related to clinical competence. Thus, it was decided to examine how well the two subsets of grading devices (i.e., concepts and principles measures and clinical skills measures) predicted (1) the variable clinical competence,

(2)

PR-

to

an

ic

to

ab

re

NE

Ta

Co

Cl

(2) the NEHPE part of clinical competence, and (3) the PR-NEHPE part of clinical competence.

Table 18 shows the concepts and principles measures to be significantly related to clinical competence, NEHPE, and PR-NEHPE while the clinical skills measures are significantly related to clinical competence and NEHPE, but not to PR-NEHPE. The concepts and principles measures correlate about twice as well with NEHPE and PR-NEHPE (.4624 and .2986, respectively) as the clinical skills measures correlate with NEHPE and PR-NEHPE (.2652 and .1441, respectively).

Table 18. Grading devices as predictors of clinical competence

Subset of Grading Devices	Dependent Variable: Clinical Competence (NEHPE + PR-NEHPE)	NEHPE	PR-NEHPE
Concepts and principles measures	$p < .0001$*	$r = .4624$ $p < .0001$*	$r = .2986$ $p < .0050$*
Clinical skills measures	$p < .0316$*	$r = .2652$ $p < .0131$*	$r = .1441$ $p < .1831$

*Significant.

pr
cl
cl
cl
(e
ag
be
a
cl
me
the
in
di
ph
co
su

These results indicate that the concepts and principles measures and clinical competence are more closely related than the clinical skills measures and clinical competence. This is indeed surprising, for the clinical skills measures require similar decisions as NEHPE (e.g., making a diagnosis, ordering laboratory tests, managing and treating the patient). Three explanations should be considered: (1) three clinical skills measures are not a sufficient number of "items" for a reliable and valid clinical skills test; (2) the individual clinical skills measures are not reliable measures in themselves; and (3) the vicarious nature of decision-making which took place in the TV Case Evaluations and the PMP is sufficiently different from the "hands-on" neurological history and physical examination of a simulated patient that the correlation between NEHPE and the clinical skills measures suffers drastically.

be

di

Th
Co

wa

ic

in

a

Ma

in

co

pr

gr

th

CHAPTER V

DISCUSSION OF THE RESEARCH RESULTS AND THEIR IMPLICATIONS FOR FUTURE RESEARCH AND PRACTICE

Research Results and Their Implications for Future Research

The research results for the three hypotheses will be reported and their implications for future research discussed.

The Relationship Between Clinical Competence and Grades

A principal motive for conducting the current study was to determine the relationship between grades and clinical competence with the intent of providing substantive information to those educators interested in establishing a competency-based instructional model for medical education. Many advocates of competency-based instruction would eliminate grades claiming that they are unrelated to clinical competence (see Wingard and Williamson, 1973). However, previous studies have investigated the relationship between grades and clinical competence after students had completed the classroom-laboratory portion of their medical training.

t
(p
bo
le
st

th
te
fe
ur
ti
wh
re
ne
ca
vi
mo
ed

be
an
am

The current study investigated the relationship between grades and clinical competence while the student-physician was in the classroom-laboratory portion of his medical training and found the two variables to be related ($p < .0001$). Thus, evidence is offered for the use of both clinical competence measures and grading devices (at least as utilized in the Neuromuscular Instructional System studied) within the competency-based instruction framework.

It is recommended that the hypothesis dealing with the relationship between clinical competence and grades be tested in other medical education courses dealing with different body systems (i.e., cardiovascular, respiratory, urinary). This replication research would probe the question of domain-specificity. That is, does a relationship which holds for one area of medical education (e.g., the relationship between clinical competence and grades in the neuromuscular system) also hold for other areas (e.g., the cardiovascular system)? Domain-specificity research is vital to the development of a competency-based instructional model whose application would extend to a variety of medical education courses.

Further research investigating the relationship between clinical skills measures (e.g., TV case evaluations and Patient Management Problems) and clinical competence among medical students in the classroom-laboratory portion

o
s
m
p

s
t
s
g
f
F
p
w
m
t
s

c
a
e
t
a
a
c
d

of their training should be undertaken. This research should focus on the optimal number of clinical skills measures (clinical cases) required to achieve satisfactory prediction in regard to clinical competence.

Even if future related research reveals clinical skills measures not to correlate well with clinical competence measures and other grading devices, medical educators should be hesitant to eliminate clinical skills measures as grading devices. Clinical skills measures may be of value for reasons other than their correlation with other measures. For example, it may be that while measures of concepts and principles correlate better than clinical skills measures with clinical competence, clinical skills measures may mediate this relationship by allowing the medical student to integrate the concepts and principles in a practical situation.

As for the lack of correlation between grades and clinical competence in previous studies, it is possible, as Wingard and Williamson (1973) suggest, that intervening experiences (e.g., internships, residencies, the effects that the demands of medical practice have on habits, attitudes, and interpersonal skills) between medical school and the time when the measures of clinical competence (or career performance) have been gathered significantly distorted any actual relationship. The current study e

a

t

e

i

q

t

F

w

l

w

p

f

s

i

m

c

d

r

a

as well as the carefully conceived and conducted research of Peterson et al. (1956) supports this hypothesis. Future research investigating the "intervening experiences hypothesis" would find the replication research described above as vital input.

In addition, survey research could be conducted to obtain physician perceptions of the impact of intervening experiences on clinical competence. The results of this initial survey research might lead to an experimental or quasi-experimental design which "teases out" complex relationships and possible explanations for these relationships. For example, one such research design of potential merit would be a two-factor design in which GPA with a number of levels was one factor and type of professional practice with levels of interest such as general practitioner, pediatrician, neurologist, orthopedic surgeon was a second factor. With clinical competence as the dependent variable, such a design might yield interesting interaction findings in which less than the highest GPA level would provide the most competent clinicians within a type of medical practice.

The question of the relationship between grades and clinical competence is a critical one. Many important decisions are based on the assumption that grades are a reflection of a student-physician's ability to perform in a clinically competent manner. Medical students are

p
T
i
p
I
e
a
i
g
s
p
a

l
b
e
t
r
s
m
I
e
c

promoted or eliminated from school on this assumption. Their potential and accomplishment are noted at points in their student career by using grades as a standard. Remediation is undertaken on the basis of grading results. Internship, residency, and fellowship selection is influenced heavily by grades. To some degree medical schools assess the achievement of their educational goals and initiate modification of those goals on the basis of student grades. Finally, as has been emphasized in the current study, the relationship between grades and clinical competence has important implications for the development of a competency-based instructional model for medical education.

The Relationship of Clinical Competence and Grades to Entry Level

Also of interest was the relationship between entry levels and clinical competence. Entry levels were found to be related to clinical competence ($p < .0270$) suggesting that entry levels such as cognitive medical knowledge, history-taking skill, physical examination skill, and doctor-patient relationship skill do have potential as pretest devices for screening, placement, and individualized instruction in a medical education course. In regard to the Neuromuscular Instructional System, history-taking skill and physical examination skill were found to be the only predictors of clinical competence.

tc

cu

th

al

va

ar

hy

or

vi

re

Ur

pr

Th

wi

bi

ar

is

sc

gr

re

at

es

Finally, entry levels were found to be related to grades ($p < .0044$). While not of vital interest in the current study, the evaluation of this hypothesis concerning the relationship between these two variables completed all meaningful regression equations involving the three variables under study (i.e., clinical competence, grades, and entry levels). In addition, the evaluation of this hypothesis provides evidence to the medical educator intent on using pre-measures for screening, placement, and individualized instruction purposes when grading devices represent the criterion measures.

Githens et al. (1970), in their study of 88 entering University of Colorado medical students, offer support for premeasuring as a placement procedure in medical education. These researchers found that entering medical students vary widely in their basic sciences preparation (i.e., anatomy, biochemistry, and physiology) and thus advanced standing and individualized placement may be possible.

The logical extension of this entry level research is future studies which investigate what entry level score(s) assure(s) minimal clinical competence or a passing grade. The current study suggests the feasibility of such research. To be most meaningful, a criterion level indicative of satisfactory clinical competence would need to be established.

hyp

tha

pen

(DP

equ

Not

cli

ski

dro

whe

dep

inv

the

nic

we

so

pa

co

an

of

an

An

dr

in

ca

A replication study involving the entry level hypotheses is advised. In Hypothesis 2 it was unexpected that one of the clinical competence entry levels (independent variables)--doctor-patient relationship skill (DPR-PD)--would drop out of the most parsimonious regression equation when clinical competence was the dependent variable. Not quite as surprising but still of interest is why the clinical competence entry levels of HIST-PD (history-taking skill), PE-PD (physical examination skill), and DPR-PD dropped out of the most parsimonious regression equation when entered as independent variables with grades as the dependent variable. Research needs to be undertaken which investigates at least three alternative explanations for these entry levels dropping out of specific most parsimonious regression equations: (1) HIST-PD, PE-PD, and DPR-PD were unreliably measured, (2) HIST-PD, PE-PD, and DPR-PD are somewhat different for a physical diagnosis course as compared to a neuromuscular system, or (3) the early data collection of HIST-PD, PE-PD, and DPR-PD in late February and early March may not have been a meaningful measurement of these skills to compare to similar evaluation in July and August for a novice population of student-physicians. Any statistical explanation for these three entry levels dropping out is not credible in that each entry level shown insignificant as reported in this study was also insignificant when entered first into a specific regression equation.

R
P

p

c

e

s

t

l

Replication Research Involving Patient Ratings

In both Hypothesis 1 and Hypothesis 2 the simulated patient's ratings of the student's ability to perform a complete neurological evaluation history and physical examination (PR-NEHPE) dropped out of the regression equation. This occurred because with the student's ability to perform a complete neurological evaluation history and physical examination (NEHPE) controlled (i.e., entered first in the regression equation), the independent variable(s) predicted a similar "part" of PR-NEHPE as was already predicted in NEHPE.

However, it is questionable that PR-NEHPE was accurately measured. The correlation between the student-physician's first and second simulated patient rating was only .24. While this correlation may have been low because the student-physician varied in his communication and interpersonal skills from case to case depending on his ability to diagnose and manage a particular case, another explanation is also possible. In order to maximize the instructional impact, student-physicians received the patient ratings of their complete neurological evaluation history and physical examinations (PR-NEHPE). Knowing this, the simulated patient probably tended to be less discriminating (i.e., less critical) than if the simulated patient rating

for
stu
whi
reg
pat
con

re
on
st

Th
Co

to
co
by
in
of
Sy
be
wo
ra
to

form (i.e., the PR-NEHPE measure) was not returned to the student-physician. Replication research is recommended which would investigate if PR-NEHPE drops out of the regression equation when the student-physician and simulated patient know that the simulated patient rating form is confidential.

Practical Implications Resulting from the Study

The practical implications resulting from the research results will be discussed followed by a report on three additional benefits which emanated from the current study.

The Relationship Between Clinical Competence and Grades

The current study concluded that grades are related to clinical competence. Thus, the development of a competency-based instructional model would be promoted by follow-up evaluative research and practice which innovated further in the utilization of a wide variety of grading devices as in the Neuromuscular Instructional System of the current study. Hopefully, higher correlations between clinical competence and specific grading devices would be obtained. In the current study correlations ranging from .182 to .392 for individual measures used to obtain grades (e.g., Neurosciences Final Test, Patient

Ma
T-
and
19
wou
con
ref
ins

var
ski
ind
ica
and
abi
his
.18
stu
eva
.107
thre
vari
anal
posi
to .

Management Problem, TV Case Evaluation I) with the mean T-score for two complete neurological evaluation history and physical examinations (NEHPE) were obtained (see Table 19). More careful development of tests and other measures would not only result in higher correlations with clinical competence but also be a step toward the criterion-referenced tests necessary for the competency-based instructional model.

The relationship between the two clinical competence variables and the grading devices which measure clinical skills should be submitted to close scrutiny. When taken individually the three grading devices which measure clinical skills--TV Case Evaluation I, TV Case Evaluation II, and the Patient Management Problem--correlate with the ability to perform a complete neurological evaluation history and physical examination (NEHPE) .185, .243, and .182, respectively, and with the patient ratings of the student's ability to perform a complete neurological evaluation history and physical examination (PR-NEHPE) .107, .132, and .060, respectively. However, when the three clinical skills measures are combined into one variable, as in the section of Chapter IV on additional analysis of the data, the correlation between this composite measure of clinical skills and (1) NEHPE increases to .265 and (2) PR-NEHPE increases to .144.

Tab

NEH
FR-

Table 19. Correlation between measures used to obtain grades and measures of clinical competence[a]

	NAMID	NSMID	NAFIN	NSFIN	NAPRAC	TVCSI	TVCSII	TVPMP
NEHPE	.341	.310	.385	.392	.319	.185	.243	.182
PR-NEHPE	.250	.248	.288	.229	.123	.107	.132	.060

[a]Symbols for Table 19.

Measures used to obtain grades:
NAMID--Neuroanatomy Mid-Term Test
NSMID--Neurosciences Mid-Term Test
NAFIN--Neuroanatomy Final Test
NSFIN--Neurosciences Final Test
NAPRAC--Practical Neuroanatomy Test
TVCSI--TV Case Evaluation I
TVCSII--TV Case Evaluation II
TVPMP--Televised Patient Management Problem

Measures of clinical competence:
NEHPE--Measure of the ability to perform a complete neurological evaluation history and physical examination.
PR-NEHPE--Patient rating of student in performing a complete neurological evaluation history and physical examination.

the
is
abi
one
med
mea
tha
sho
and
is
var
the
of

reg
mea
a s
nov
are
cor
cli
pat
gra
ind

The increase in correlation resulting from combining the three clinical skills measures into a composite measure is analogous to the increment in reliability and predictability which emanates from adding items to a test. Thus, one of the practical findings from this study is that medical educators wishing to assess clinical skills by means of presenting specific cases are advised to give more than just one or two cases. The data presented here have shown that the correlation between clinical skill measures and clinical competence increases when a composite measure is formed from three individual cases. Depending on a variety of circumstances (e.g., the subject matter content, the homogeneity of the student group), the optimal number of cases for reliable and predictive measurement will vary.

In addition, armed with a variety of valid (in regard to predicting clinical competence) and reliable measures, the medical educator could report much more than a single overall grade for his particular course. He could now evaluate and grade students in a number of different areas--e.g., knowledge and understanding of fundamental concepts and principles, application of knowledge to the clinical case, ability to communicate and relate to the patient. This more specific and meaningful evaluation and grading would be of greater value in making decisions about individual students and medical education programs.

The Relationship of Clinical Competence and Grades to Entry Levels

Consistent with a recommendation issued by MSU-COM's Ad Hoc Committee on Evaluation Procedures (1975), evaluative research and practice into the predictive value of various entry level skills (cognitive, psychomotor, and affective) for particular body systems (e.g., cardiovascular, urinary) could be conducted with the purpose of developing procedures for screening, placement, and individualized instruction consistent with the competency-based instructional model. For example, data could be gathered on the relationship of cognitive knowledge, psychomotor skill, history-taking skill, physical examination skill, and doctor-patient relationship skill to grades and clinical competence in such body systems as the cardiovascular and urinary systems. The results might be quite different from those reported for the Neuromuscular Instructional System. Such a research program would permit screening, placement, and individualized instruction within specialized medical education courses.

Three Additional Benefits of the Study

In order to perform the study herein reported, it was necessary to carefully coordinate research design with instruction and evaluation in the Neuromuscular Instructional System. In the course of coordinating the research, valuable insights were gained in the use of physician's assistants and simulated patients. Also, the simulated patient rating form (see Appendix C) used by the simulated patient in evaluating student-physician ability to perform a complete neurological evaluation history and physical examination was refined and evaluated with implications for its future use. These three practical benefits emanating from the current study will be described below.

Utilization of physician's assistants.--Developed by Stead (1966) at Duke University in 1965, physician's assistants (PA) were seen as providing much needed assistance to overworked general practitioners. In fact, the PA has proven to be a partial solution to the health care manpower shortage in the areas of primary care and preventive medicine as well as an asset to specialty and emergency medicine. In addition, the Neuromuscular Instructional System used PAs in an innovative manner to facilitate the accomplishment of the objectives of the "hands-on" clinical component of the instructional system. These objectives focus on training students to (1) perform the basic psychomotor skills necessary in a neurological evaluation of a patient and (2) make appropriate diagnostic decisions on the basis of the findings of a neurological evaluation. Adequate guidance during the clinical component of the system was made possible by the utilization of PAs, who functioned in four areas: (1) supervision of "self-instructional" laboratories, (2) screening of psychomotor skills, (3) training of simulated patients, and (4) provision of follow-up discussions of simulated patient cases. Two PAs were hired to work twenty hours per week each for a period of ten weeks.

During the first four weeks, students completed five individually-paced "self-instructional" laboratory units which incorporated the basic psychomotor skills and conceptual knowledge needed for performing a neurological evaluation. PAs were available and readily used for individual consultation in regard to proper techniques and procedures. During the fifth week and before any student-simulated patient contact, all students were required to perform an extensive neurological psychomotor practical test on a fellow student in the presence of a PA, who scored this performance. At the conclusion of the test, the student was given feedback on his performance. Six students scored below the established standard of competence and were required to undergo remediation by the PAs.

During the 10-week Neuromuscular Instructional System, PAs were involved in the training of simulated patients. The Neuromuscular Instructional System coordinator, after selecting the cases to be simulated and doing the initial programming of patients, instructed the PAs in ways and means of checking simulated patients during their second training session to assure a high quality of patient reality and consistency. The PAs met with simulated patients individually for this second training session (one-half hour in duration), and, in addition, were available for "on-the-spot" checks before student-patient interaction, which took place during weeks six through nine as each first-year student performed two complete neurological evaluation history and physical examinations. PAs also met with small groups of students immediately after they had completed the write-up of a neurological case to discuss the case in detail.

In addition to their utilization for instructional purposes, the PAs were key elements in the completion of the current study. First, the neurological psychomotor practical test administered by the PAs was used in the additional analysis of the data reported in Chapter IV. Second, data for one of the clinical competence variables, the two complete neurological evaluation history and physical examinations with simulated patients performed by each student, could not have been gathered without the PAs in that they were responsible for (1) the training and standardization of simulated patient performance and (2) scoring the examinations using criteria in large part determined by the PAs.

In conclusion, both the Neuromuscular Instructional System coordinator and the first-year osteopathic medical students agreed that the PAs added a valuable instructional dimension to the clinical component of the system. Ninety-seven percent of the students stated that the PAs were helpful or very helpful to their learning. Through their (1) instruction in and screening of the necessary neurological psychomotor skills, (2) training of simulated patients, and (3) small group discussions dealing with a specific neurological case immediately after students had performed a complete neurological evaluation history and physical examination with a simulated patient programmed for the case, both the coordinator and the students felt that the students were better able to complete neurological evaluations. In addition, PAs were essential in the gathering of data needed for the current study.

Finally, the use of PAs in the described manner would, in all probability, receive a strong endorsement from a cost-effective analysis. Two PAs, each working 20 hours per week, were able to competently perform tasks which otherwise clinical faculty would have been called upon to perform--i.e., supervise "self-instructional" laboratories, screen psychomotor skills, train simulated patients, provide follow-up discussions of simulated patient cases, and grade neurological evaluation history and physical examinations. At slightly more than one-third the cost ($2,700 versus $7,200), the PAs provided a service at the same high level of quality as could have been provided with clinical faculty.

Physician's assistants will be utilized in future medical education courses at MSU-COM in a manner analogous to that described here.

Utilization of simulated patients.--The system coordinator, the researcher, and the physician's assistants gained valuable insights into the selection and training of simulated patients as well as the logistical problems of coordinating such a large program. In a four-week period, 89 students performed 178 complete neurological evaluation history and physical examinations and had the opportunity to discuss in detail the case with a trained professional immediately after they had written their report on the case. To implement such an extensive program, 19 simulated patients had to be programmed to simulate either one or two neurological disorders.

From the experience of implementing the simulated patient program for the Neuromuscular Instructional System, the system coordinator, researcher, and physician's assistants gathered valuable educational and administrative information which can be used in subsequent simulated patient programs. Some examples of insights emanating from the Neuromuscular Instructional System simulated patient program follow.

Simulated patients can typically be trained to simulate neurological disorders by the procedure described below:

1. a one-half hour small group meeting with a trained medical professional as trainer and four to eight simulated patients at which the case(s) is (are) described and initial programming begun;

2. a 20 to 30 minute follow-up session in which a trained medical professional as trainer works with each simulated patient on a one-to-one basis to perfect the behavior and reactions of the simulated patient;

3. on-the-spot checks before the simulated patient performs the role for the first time and at subsequent performances as needed.

Student-physicians and simulated patients will respond with punctuality and enthusiasm if they see their efforts as being worthwhile. For the student-physicians only 3 of 178 appointments were not begun as scheduled, and for simulated patients only 2 of 178 appointments were not kept as scheduled. Student-physicians and simulated patients were very tolerant of any logistical problems which occurred during the three-hour afternoon sessions. Almost always both student-physicians and simulated patients displayed positive affect following the appointment and commented on the value of the experience. Table 20 shows that student-physicians, when asked to rate the simulated

Table 20. Student-physician rating of simulated patient examination experience

Description of Experience	Mean Rating[a]	Description of Your Feelings About Yourself	Mean Rating[a]
Useful	6.30	Pleasant	5.93
Meaningful	6.12	Successful	5.89
Important	6.10	Secure	5.81
Successful	5.78	Calm	5.75
Realistic	5.02	Competent	5.69

[a]Ratings were based on a 7-point semantic differential scale with 7 being the highest rating.

patient experience, had high mean ratings on a 7-point semantic differential scale (see Appendix G). Table 20 also reveals that, when asked to describe "your feelings about yourself" after completing the simulated patient examination, student-physician ratings yielded high mean values.

Simulated patients must be given extensive practice and training in the rating of student-physicians using the simulated patient rating form. The researcher speculated that PR-NEHPE (the patient ratings of the student's ability to perform a complete neurological evaluation history and physical examination) dropped out of the most parsimonious regression equation for both Hypothesis 1 and Hypothesis 2 because of unreliable rating by the simulated patients.

In order to maximize the instructional impact, student-physicians received the patient ratings of their complete neurological evaluation history and physical examination (PR-NEHPE). Knowing this, the simulated patient probably tended to be less discriminating (i.e., less critical) than if the simulated patient rating form was not returned to the student-physician. Future evaluative research and practice using the simulated patient rating form as a criterion measure should use one of two alternative procedures. First, if the student-physician is to be given the results of the patient ratings, simulated patients should be better sensitized to the need for critical, discriminating ratings. Second, at the loss of valuable feedback to student-physicians, the simulated patient rating forms could not be returned to the student-physicians.

Like the employment of PAs in the Neuromuscular Instructional System, the utilization of simulated patients proved to be a cost-effective procedure. The total cost of the simulated patient component (19 simulated patients doing 178 complete neurological evaluation history and physical examinations) was $896.46.

The insights herein described have been and will be valuable to MSU-COM staff as they continue to employ simulated patients as a vital instructional resource in clinical education.

Simulated patient rating form.--The simulated patient rating form (see Appendix C) was developed from Tinning (1973), who used the Very Interested-Very Uninterested scale, the Secure-Insecure scale, and the Very Gentle During the Physical Exam-Very Rough During the Physical Exam scale and Hess (1969) and Turner et al. (1972), who used the Frequently Used Uncommon Words or Concepts-Did Not Use Any Uncommon Words or Concepts scale, the Clear-Unclear scale, the Very Understanding-Not Understanding scale, and the Patient-Impatient scale. The coefficient alpha reliability index for internal consistency was .82033 indicating that the simulated patient rating form contained scales which were measuring generally the same variable--i.e., doctor-patient relationship skill. Intercorrelations among the scales ranged from .2352 to .6419, and the coefficient alpha with one of the seven scales deleted never dropped lower than .77682 (see Table 21).

The statistics reported above is evidence for the usefulness of the simulated patient rating form used in the current study as a measure of doctor-patient relationship skill. Correspondingly, the simulated patient rating form could be utilized in any medical education course involving a clinical evaluation component. MSU-COM's Ad Hoc Committee on Evaluation Procedures (1975) has urged such usage.

Table 21. Simulated patient rating form--intercorrelation matrix for scales[a]

	INT	SEC	GENT	WORD	CLAR	UND	PAT
INT	1.000						
SEC	.2772	1.000					
GENT	.4591	.3262	1.000				
WORD	.2705	.2371	.3106	1.000			
CLAR	.5592	.3896	.5231	.2357	1.000		
UND	.6419	.4275	.5539	.2352	.5107	1.000	
PAT	.5194	.2523	.5522	.3961	.3200	.3974	1.000

Coefficient Alpha with Scale Deleted

INT	.77929
SEC	.81572
GENT	.78432
WORD	.82680
CLAR	.78925
UND	.77682
PAT	.79717

[a]Symbols used for Table 21.

INT--Very Interested-Very Uninterested scale
SEC--Secure-Insecure scale
GENT--Very Gentle During the Physical Exam-Very Rough During the Physical Exam scale
WORD--Frequently Used Uncommon Words or Concepts-Did Not Use Any Uncommon Words or Concepts scale
CLAR--Clear-Unclear scale
UND--Very Understanding-Not Understanding scale
PAT--Patient-Impatient scale.

It is also possible that the simulated patient rating form or a similar doctor-patient relationship measure could be used for those board certification examinations or state and federal licensing examinations which assess clinical competence.

Summary

This chapter discussed the research results of the current study in the context of their implications for future research and practice in medical education in general and competency-based instruction in particular. A number of considerations were presented. First, the feasibility of using both grading devices and clinical competence measures within the competency-based instruction framework was suggested. Second, replication research investigating the relationship between clinical competence and grades in other areas of medical education was recommended. Third, the value of further investigation into the relationship between clinical skills measures and clinical competence was pointed out. Fourth, it was urged that the "intervening experiences hypothesis" be studied systematically. Fifth, it was recommended that research into entry level variables appropriate for screening, placement, and individualized instruction within other specialized medical education courses be conducted. Sixth, replication research designed to investigate the elimination of PR-NEHPE and the clinical competence entry levels of HIST-PD, PE-PD, and DPR-PD from various most parsimonious regression equations was suggested. Seventh, it was urged that a greater variety of grading devices be introduced into medical education courses and that improving their reliability and validity be the concern of measurement specialists. This broader approach to grading could be a step toward criterion-referenced testing, an essential component of the competency-based instructional model. Eighth, the utilization of physician's assistants in the Neuromuscular Instructional System was discussed. Ninth, the educational insights gained through the implementation of the simulated patient program for the Neuromuscular Instructional System were described. Tenth, the refinement of the simulated patient rating form and its potential use in other clinical settings were discussed.

APPENDICES

APPENDIX A

TV CASE EVALUATION ANSWER SHEET

APPENDIX A

TV CASE EVALUATION ANSWER SHEET

STUDENT'S NAME ________ INSTRUCTOR'S NAME ________ COURSE ________ CASE NO. ________

NEUROLOGICAL EXAMINATION

DIAGNOSIS – LABORATORY TESTS & PROCEDURES – MANAGEMENT

DIAGNOSIS

Preliminary Diagnostic Summary

Systems Diagnosis
A 1 2 3 4 5 6 7 8 9 10 11 12

Location of Pathology
B 1 2 3 4 5 6 7 8 9 10

Categorical Neuromuscular Diagnosis
C 1 2 3 4 5 6

DIAGNOSIS

Specific Working and Differential Diagnosis Categories

Vascular (Including Cardiac)
V 1 2 3 4 5 6 7 8
9 10 11 12 13 14 15 16

Infections
I 1 2 3 4 5 6 17 18

Trauma
T 1 2 3 4 5 6 7 8

Autoimmune
A 1 2 3 4

Metabolic Problems
M 1 2 3 4 5 6 7 8

Inherited (Familial, Congenital)
I 1 2 3 4 5 9

Neoplasm
N 1 2 3 4 5 6 7 8

Skeletal and Structural
S 1 2 3 4 5 6 7

Convulsive Disorders
C 1 2 3 4 5 6 7 8

Degenerative Demyelinating
D 1 2 3 4 5 6 7 8

LABORATORY TESTS AND PROCEDURES

Specific Neurodiagnostic Tests
A 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
17 18 19 20 21 22 23

Blood and Serum
B 1 2 3 4 5 6 7 8 9
10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25
26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41
42 43 44 45 46 47 48 49

Radioisotope
C 1 2 3 4 5 6 7

Stool
D 1 2 3

Sputum
E 1 2 3

Urine
F 1 2 3 4 5 6 7 8

Miscellaneous Tests
H 1 2 3 4 5 6 7 8 9 10 11 12

X-rays
I 1 2 3 4 5 6 7 8 9 10

MANAGEMENT

Pharmacologics
A 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
16 17 18 19 20

General
B 1 2 3 4 5 6 7 8

APPENDIX B

NEUROLOGICAL EVALUATION HISTORY AND PHYSICAL EXAMINATION--THREE SIMULATED PATIENT CASES

APPENDIX B

NEUROLOGICAL EVALUATION

HISTORY AND PHYSICAL EXAMINATION

Simulated Patient Case No. 1

DATE: *Summer 1975* TYPE OF CASE: *Neurological*

STUDENT: *Jane Doe* PATIENT'S AGE: *22*

INSTRUCTOR: *Dr. Jacobson* RACE: *Caucasian/Female*

1. CHIEF COMPLAINT(S)--(3 points): *Neck pains with some radiation down right arm and associated numbness and weakness--approximately three months duration.*

2. ONSET AND COURSE OF CHIEF COMPLAINT(S)--(6 points):

The patient was seen in the office noting neck pain with some distinct radiation down the right arm as well as some associated numbness--most evidenced in the thumb and index finger of the right hand and some minimal weakness in right-hand function. This discomfort apparently began with an automobile accident three months ago at which time the patient's car was struck in the rear, and the patient apparently suffered "whiplash" injury. In addition, the patient's head struck the front windshield, and she suffered some severe contusions of the forehead. The patient also reports some dizziness and cephalgia, which were present at the onset of this discomfort, but appear to have been resolved during the last three months, but still cause periodic discomfort. The patient also notes a good deal of associated anxiety and tension since her accident. Analgesic medication has given little relief from her discomfort, and the patient apparently is also being treated with tranquilizers prescribed by a physician who has been examining her since the accident.

3. PAST HISTORY--(5 points)
FAMILY HISTORY (Mother, Father, Wife, Siblings, Children)

Endocrine Dysfunction	*Diabetes (mother)*
Cancer	____________
Tuberculosis	____________
Neurosis, Psychosis	*Epilepsy (sister)*
Cardiovascular Disease	*Hypertension (mother)*
Other	____________

MEDICAL HISTORY:

Previous Hospitalization: *Gastrointestinal infection with genito-urinary complications (1974).*

Allergies: *None*

Medications: *Valium*
Darvon compound

Accidents: *Car accident (3 months ago)*

Surgery: *Right inguinal herniorraphy (1973)*

Diseases:

Habits: *Smokes (1 package per day for 5 years).*

SOCIAL HISTORY (Work, Hobbies, Recreation):
Student (nutrition); waitress
Snow skier
Gourmet cook

4. SYSTEMS REVIEW (if appropriate) (Gynecological, Obstetrical, Gastrointestinal, Genito-urinary, Cardiovascular, Respiratory, Metabolic, Neuro-Muscular)--(2 points):

Neuromuscular--none

5. PHYSICAL EXAMINATION--(17¼ points):

GENERAL APPEARANCE:

Coloration (Skin, Sclera): *Good*

Physical Development (Asthenic, Obese, etc.)

Asthenic

GENERAL FINDINGS:

BP: *98/70*

Cardiac Auscultation

Rate *68*

Rhythm *regular*

Murmurs *normal*

Neck Auscultation

Bruits *none*

Ophthalmoscopic (GRI-IV)

Vessels *normal*

Disc *normal*

Retina *normal*

MENTAL STATUS:

State of Consciousness (check)

X Alert

_____ Unconscious or comatose

_____ Confused or obtunded

_____ Decerebrate or decorticate

Speech and Language Function

N Aphasic or dysphasic

N Dysarthric or anarthric

REFLEXES:

Deep Tendon
(Designate 0-Absent, 1-Hypoactive, 2-Normal, 3-Hyperactive)

	Left	Right
Patellar	2	2
Biceps	2	2
Tricep	2	2
Brachioradial	3	2
Achilles	2	2

Pathological or Superficial
(Indicate A-Absent, P-Present, E-Equivocal)

	Left	Right
Plantar	P	A
Babinski	A	P
Ankle Clonus	A	A
Abdominal	P	P
Hoffman	A	P

SENSORY:
(Indicate A-Abnormal: Hypoactive or Hyperactive Responsive, N-Normal, E-Equivocal Dysfunction)

	Response	Location
Vibration	N	Rt. big toe
Pinprick	A	Rt. thumb &
Light touch	A	index finger
Position sense	N	Rt. big toe
Stereognosis	N	Rt. hand

MUSCLE FUNCTION AND GAIT: (check appropriate headings)

Fasiculations Yes _____ No __*X*__

Gait: Normal __*X*__

Abnormal _____ (describe also)

____

Muscle Tone:

Spastic _____ Flaccid _____

Rigid _____ Normal __*X*__

Muscle Strength (indicate specific muscle weakness)

Weakness in right biceps, brachioradialis,

and grasp.

____

____

CEREBELLAR AND DORSAL COLUMN FUNCTIONS (Indicate N-Normal, A-Abnormal, E-Equivocal)

__*N*__ Finger to nose

__*N*__ Dysdiadochokinesia

__*N*__ Tandem gait

__*N*__ Heel to knee

__*N*__ Romberg

EXTRAPYRAMIDAL (check appropriate headings)

__*N*__ Spontaneous movements (describe)

__*N*__ Cog wheel rigidity

__*N*__ Mask like facies

__*N*__ Decreased eyeblinks

__*N*__ Loss of arm swing

CRANIAL NERVES:
(Indicate nerves checked and if pathology present)

1. *All within normal limits*
2. "
3. "
4. "
5. "
6. "
7. "
8. "
9. "
10. "
11. "
12. "

6. SUMMARY--($4\frac{1}{4}$ points)

a. General Results: (check)

_____ Normal Neurological

__*X*__ Abnormal Neurological

_____ Equivocal

b. Assessment of Area of Neurological System Dysfunction (check)

__*X*__ Motor

__*X*__ Sensory

_____ Mentation and Behavior

c. Anatomical Location (check)

_____ Primary Muscle Dysfunction

__*X*__ Peripheral Nerve or Root Dysfunction

_____ Spinal Cord

_____ Brain Stem

_____ Cerebral Hemispheres

7. PROVISIONAL OR WORKING DIAGNOSIS(ES)
(Include all systems)--(1 point)

Categorical Diagnosis (neuropathy, encephal-opathy, etc.)	Specific (tumor, cerebral hemorrhage, etc.)
Radioculopathy	*Cervical Disc Herniation*

8. DIFFERENTIAL DIAGNOSIS (Identify no more than three neurological disorders)--(6 points):

General Pathology Category	Specific Type of Pathology
Vascular	*Thoracic Outlet Syndrome*
Infectious	
Traumatic	*Ruptured Disc*
Autoimmune	
Metabolic	*Diabetes*
Inherited	
Neoplastic (or mass lesion)	*Tumor*
Cardiac Dysfunction	
Degenerative or Demyelinating	
Others	*1. Cervical Spondylosis and/or Osteoarthritis*
	2. Carpal Tunnel

9. TESTS (laboratory tests and other diagnostic procedures)--(4 points):

Specific Neurodiagnostic Tests (EEG, lumbar puncture, etc.)	General Laboratory Tests (CBC, urinalysis, etc.)
Cervical Spine X-Ray	*Chest X-Ray*
	CBC, ESR, RBS, 2 Hr. PPBS
EMG	*GTT, UA*
Nerve Conduction	
Lumbar puncture and myelogram	

10. THERAPY--(2 points):

Specific	Supportive
Cervical Traction	*Analgesics*
Surgery for cervical decompression	*Muscle relaxant*
(Laminectomy with fusion)	*Heat (physical therapy)*

COMMENTS:

Note: All practical Exams have maximum score of 50.5 points.
All Exams are standardized on a 50.5 point basis.

NEUROLOGICAL EVALUATION

HISTORY AND PHYSICAL EXAMINATION

Simulated Patient Case No. 2

DATE: *Summer 1975* TYPE OF CASE: *Neurological*

STUDENT: *John Doe* PATIENT'S AGE: *24*

INSTRUCTOR: *Dr. Jacobson* RACE: *Caucasian/Male*

1. CHIEF COMPLAINT(S)--(3 points): *Back pain (six weeks duration) with right leg radiation.*

2. ONSET AND COURSE OF CHIEF COMPLAINT(S)--(6 points): *The patient was seen because of severe back pain for the past six weeks. Pain evidently was initiated while the patient was performing some stretching and lifting maneuvers at home at which time he feels he greatly "strained" his back. The patient also has noted some numbness and tingling in the posterior, lateral thigh extending down lateral leg to big toe with some episodic pain radiation that tends to be noted in the same area. The patient has found that rest seems to help his discomfort and is greatly aggravated by activity and exercise. Heat, hot baths, and Ben-Gay have all afforded negligible relief. The patient has also utilized aspirin and other common analgesics without success. He reports, in addition, having seen a chiropractor several days following the onset of discomfort but noted no distinct relief following manipulative procedures.*

3. PAST HISTORY--(5 points)
FAMILY HISTORY (Mother, Father, Wife, Siblings, Children)

Endocrine Dysfunction	
Cancer	
Tuberculosis	
Neurosis, Psychosis	
Cardiovascular Disease	*HBP (father)*
Other	

MEDICAL HISTORY:

Previous Hospitalization: *T & A (1953)*
Rheumatic Fever (1953)

Allergies: *None*

Medications: *Aspirin*
Excedrin
Ben-Gay

Accidents: *None*

Surgery: *T & A (1953)*

Diseases: *Childhood Diseases: German Measles*
Mumps
Chicken Pox

Habits: *No Smoking*
No Drinking

SOCIAL HISTORY (Work, Hobbies, Recreation):

Student, tennis, jogging

4. SYSTEMS REVIEW (if appropriate) (Gynecological, Obstetrical, Gastrointestinal, Genito-Urinary, Neuro-Muscular)--(2 points):

Neuromuscular--none

5. PHYSICAL EXAMINATION (17¾ points):

GENERAL APPEARANCE:

Coloration (Skin, Sclera): *No abnormal pigmentation*

Physical Development (Asthenic, Obese, etc.): *Medium Build*

GENERAL FINDINGS:

BP: *108/80*

Cardiac Auscultation

Rate *76*

Rhythm *regular*

Murmurs *none*

Neck Auscultation

Bruits *negative*

Ophthalmoscopic (GRI-IV)

Vessels *no A-V nicking*

Disc *well-delineated*

Retina *no pigmented areas*

MENTAL STATUS:

State of Consciousness (check)

X Alert

_____ Unconscious or comatose

_____ Confused or obtunded

_____ Decerebrate or decorticate

Speech and Language Function

N Aphasic or dysphasic

N Dysarthric or anarthric

REFLEXES:

Deep Tendon
(Designate 0-Absent, 1-Hypoactive, 2-Normal, 3-Hyperactive)

	Left	Right
Patellar	*2*	*2*
Biceps	*2*	*2*
Tricep	*1-2*	*1-2*
Brachioradial	*2*	*2*
Achilles	*2*	*2*

Pathological or Superficial
(Indicate A-Absent, P-Present, E-Equivocal)

	Left	Right
Plantar	*P*	*P*
Babinski	*A*	*A*
Ankle Clonus	*A*	*A*
Abdominal	*P*	*P*
Hoffman	*A*	*A*

SENSORY:
(Indicate A-Abnormal: Hypoactive or Hyperactive Responsive, N-Normal, E-Equivocal Dysfunction)

	Response	Location
Vibration	*N*	*Rt. big toe*
Pinprick	*A*	*Lateral side of right leg*
Light touch	*A*	*& dorsum foot & large toe*
Position sense	*N*	*Rt. big toe*
Stereognosis	*N*	*Rt. hand*

MUSCLE FUNCTION AND GAIT: (check appropriate headings)

Fasiculations Yes _____ No _X_

Gait: Normal _____

Abnormal _X_ (describe also)

Slightly slow and guarded (Antalgic) slight limp right leg

Muscle Tone:

Spastic _____ Flaccid _____

Rigid _____ Normal _X_

Muscle Strength (indicate specific muscle weakness)

Weakness right extensor hallicus longus and tibialis anterior

CEREBELLAR AND DORSAL COLUMN FUNCTIONS
(Indicate N-Normal, A-Abnormal, E-Equivocal)

_N_ Finger to nose

_N_ Dysdiadochokinesia

_N_ Tandem gait

_N_ Heel to knee

_N_ Romberg

EXTRAPYRAMIDAL (check appropriate headings)

_N_ Spontaneous movements (describe)

_N_ Cog wheel rigidity

_N_ Mask like facies

_N_ Decreased eyeblinks

_N_ Loss of arm swing

CRANIAL NERVES:
(Indicate nerves checked and if pathology present)

1. *All within normal limits*
2. "
3. "
4. "
5. "
6. "
7. "
8. "
9. "
10. "
11. "
12. "

6. SUMMARY (4¼ points)

a. General Results: (check)

_____ Normal Neurological

__X__ Abnormal Neurological

_____ Equivocal

b. Assessment of Area of Neurological System Dysfunction: (check)

__X__ Motor

__X__ Sensory

_____ Mentation and Behavior

c. Anatomical Location: (check)

_____ Primary Muscle Dysfunction

__X__ Peripheral Nerve or Root Dysfunction

_____ Spinal Cord

_____ Brain Stem

_____ Cerebral Hemispheres

7. PROVISIONAL OR WORKING DIAGNOSIS(ES) (Include all systems)--(1 point)

Categorical Diagnosis (neuropathy, encephal- opathy, etc.)	Specific (tumor, cerebral hemorrhage, etc.)
Radiculopathy and/or	*Lumbar Disc Herniation*
Myositis	*Lumbar myositis*

8. DIFFERENTIAL DIAGNOSIS (Identify no more than three neurological disorders)--(6 points):

General Pathology Category	Specific Type of Pathology
Vascular	
Infectious	
Traumatic	
Autoimmune	
Metabolic	*Diabetes (not significant consideration)*
Inherited	
Neoplastic (or mass lesion)	*Tumor of lumbar spine*
Cardiac Dysfunction	
Degenerative or Demyelinating	
Others	*1. Spondylolithesis (anomalous skeletal conditions)*
	2. Osteoarthritis

9. TESTS (laboratory tests and other diagnostic procedures)--(4 points):

Specific Neurodiagnostic Tests (EEG, lumbar puncture, etc.)	General Laboratory Tests (CBC, urinalysis, etc.)
Lumbar Spine X-Rays	
EMG and nerve conduction studies	*CBC/FBS/2 Hr. PPBS/ESR*
Lumbar Puncture and Myelogram	

10. THERAPY--(2 points)

Specific	Supportive
Pelvic Traction	*Analgesics*
Bed rest	*Muscle relaxant*
Surgery if necessary	*Elastic or corset back support*
	Local heat
	Steroids

COMMENTS:

Note: All Practical Exams have maximum score of 50.5 points.
All Exams are standardized on a 50.5 point basis.

NEUROLOGICAL EVALUATION

HISTORY AND PHYSICAL EXAMINATION

Simulated Patient Case No. 3

DATE: *Summer 1975*

TYPE OF CASE: *Neurological*

STUDENT: *Joan Doe*

PATIENT'S AGE: *20*

INSTRUCTOR: *Dr. Jacobson*

RACE: *Caucasian/Female*

1. CHIEF COMPLAINT(S)--(3 points): *weakness right hand, memory loss, seizure episode.*

2. ONSET AND COURSE OF CHIEF COMPLAINT(S)--(6 points):

Five months ago the patient noted development of weakness of right hand over period of 24 hours. This problem has become slightly worse since that time producing occasional difficulty in use of right hand. Patient also reports that some numbness and tingling on the right side were noted about 2 months prior to the onset of weakness. During the past 2 months patient has been aware of some occasional problems with memory. Patient reports having consulted two physicians and being told she had suffered a "small stroke." Two weeks ago patient had seizure episode with loss of consciousness and shaking movements of all extremities. Taken to hospital emergency room and then discharged, she was told to see family doctor. No additional complaints are elicited other than awareness of increasing depression over present problem.

3. PAST HISTORY--(5 points):
FAMILY HISTORY (Mother, Father, Wife, Siblings, Children)

Endocrine Dysfunction	*Thyroid (father)*
Cancer	*Lung cancer (paternal grandfather)*
Tuberculosis	
Neurosis, Psychosis	
Cardiovascular Disease	
Other	

MEDICAL HISTORY:

Previous Hospitalization: *None*

Allergies: *None*

Medications: *None*

Accidents: *Broken middle right finger (1970)*

Surgery: *None*

Diseases: *Childhood diseases: measles, mumps, small pox*

Habits: *No smoking; very limited social drinking.*

SOCIAL HISTORY (Work, Hobbies, Recreation): *Student, Waitress, Guitarist, Contract Bridge Player*

4. SYSTEMS REVIEW (if appropriate) (Gynecological, Obstetrical, Gastrointestinal, Genito-Urinary, Neuro-Muscular)--(2 points):

Neuromuscular--none

5. PHYSICAL EXAMINATION (17¼ points):

GENERAL APPEARANCE:

Coloration (Skin, Sclera): *Pigmentation within normal limits*

Physical Development (Asthenic, Obese, etc.): *Slender build*

GENERAL FINDINGS:

BP: *120/76*

Cardiac Auscultation

Rate *78*

Rhythm *regular*

Murmurs *none*

Neck Auscultation

Bruits *none*

Ophthalmoscopic (GRI-IV)

Vessels *no A-V nicking*

Disc *well-delineated*

Retina *no pigmentations*

MENTAL STATUS:

State of Consciousness (check)

X Alert

_____ Unconscious or comatose

_____ Confused or obtunded

_____ Decerebrate or decorticate

Speech and Language Function

N Aphasic or dysphasic

N Dysarthric or anarthric

REFLEXES:

Deep Tendon
(Designate 0-Absent, 1-Hypoactive, 2-Normal, 3-Hyperactive)

	Left	Right
Patellar	*2*	*3*
Biceps	*2*	*3*
Tricep	*2*	*3*
Brachioradial	*2*	*3*
Achilles	*2*	*2-3*

Pathological or Superficial
(Indicate A-Absent, P-Present, E-Equivocal)

	Left	Right
Plantar	*P*	*A*
Babinski	*A*	*P*
Ankle Clonus	*A*	*A*
Abdominal	*P*	*P*
Hoffman	*A*	*P*

SENSORY:
(Indicative A-Abnormal: Hypoactive or Hyperactive Responsive, N-Normal, E-Equivocal Dysfunction)

	Response	Location
Vibration	*N*	*right big toe*
Pinprick	*N*	*arms and legs*
Light touch	*N*	*slight decrease tips of fingers right hand*
Position sense	*N*	*right big toe*
Stereognosis	*A-E*	*right hand*

MUSCLE FUNCTION AND GAIT: (check appropriate headings)

Fasiculations Yes _____ No __*X*__

Gait: Normal __*X*__

Abnormal _____ (describe also)

____

____

Muscle Tone:

Spastic _____ Flaccid __*X*__

Rigid _____ Normal _____

Muscle Strength (indicate specific muscle weakness)

Weakness right hand grasp (finger flexion)

____

____

CEREBELLAR AND DORSAL COLUMN FUNCTIONS
(Indicate N-Normal, A-Abnormal, E-Equivocal)

__*N*__ Finger to nose

__*N*__ Dysdiadochokinesia

__*N*__ Tandem gait

__*N*__ Heel to knee

__*N*__ Rombert

EXTRAPYRAMIDAL (check appropriate headings)

__*N*__ Spontaneous movements (describe)

__*N*__ Cog wheel rigidity

__*N*__ Mask like facies

__*N*__ Decreased eyeblinks

__*N*__ Loss of arm swing

CRANIAL NERVES:
(Indicate nerves checked and if pathology present)

1. *All within normal limits*
2. "
3. "
4. "
5. "
6. "
7. "
8. "
9. "
10. "
11. "
12. "

6. SUMMARY ($4\frac{1}{4}$ points)

a. General Results: (check)

_____ Normal Neurological

__*X*__ Abnormal Neurological

_____ Equivocal

b. Assessment of ARea of Neurological System Dysfunction: (check)

__*X*__ Motor

__*X*__ Sensory

__*X*__ Mentation and Behavior

c. Anatomical Location: (check)

_____ Primary Muscle Dysfunction

_____ Peripheral Nerve or Root Dysfunction

_____ Spinal Cord

_____ Brain Stem

__*X*__ Cerebral Hemispheres

7. PROVISIONAL OR WORKING DIAGNOSIS(ES) (Include all systems)--(1 point):

Categorical Diagnosis (neuropathy, encephalopathy, etc.)	Specific (tumor, cerebral hemorrhage, etc.)
Encephalopathy	*Cerebral Neoplasm*

8. DIFFERENTIAL DIAGNOSIS (Identify no more than three neurological disorders)--(6 points)

General Pathology Category	Specific Type of Pathology
Vascular	*1. Cerebral Thrombosis (middle cerebral artery or secondary to internal carotid obstruction)* *2. Vascular Malformation*
Infectious	
Traumatic	
Autoimmune	
Metabolic	*Diabetes Mellitus (predisposing to early cerebral vascular disease)*
Inherited	
Neoplastic (or mass lesion)	*1. Cerebral Neoplasm* *2. Primary Malignancy with Cerebral Metastasis*
Cardiac Dysfunction	
Degenerative or Demyelinating	
Others	

9. TESTS (laboratory tests and other diagnostic procedures)--(4 points):

Specific Neurodiagnostic Tests (EEG, lumbar puncture, etc.)	General Laboratory Tests (CBC, urinalysis, etc.)
Skull X-Rays	
Lumbar Puncture	*CVC, ESR*
EEG: Brain Scan	*FBS, 2 Hr. PPNS, GTC*
Echoencephalogram	*Chest X-Ray*
Ophthalmodynamometry	
Cerebral Angiogram and/or	
Pneumoencephalogram	

10. THERAPY--(2 points):

Specific	Supportive
Craniotomy	*Physical Therapy*
Chemotherapy (?)	*Anticonvulsants*
Radiation if appropriate	*(Dilantin and/or Phenobarbital)*
post surgically	

COMMENTS:

Note: All Practical Exams have maximum score of 50.5 points.
All Exams are standardized on a 50.5 point basis.

APPENDIX C

SIMULATED PATIENT RATING

APPENDIX C

SIMULATED PATIENT RATING

In the history and physical examination just completed, I felt that the doctor was:

1	2	3	4	5
Very Interested (paid attention to me; genuinely warm; eager to help)		Reasonably Interested (seemed interested in me and tried to help)		Very Uninterested (seemed cool and indifferent--not caring to help)
1	2	3	4	5
Frequently Used Uncommon Words or Concepts		Sometimes Used Uncommon Words or Concepts		Did Not Use Any Uncommon Words or Concepts
1	2	3	4	5
Clear (explained thoroughly what he was doing)		Average (explained adequately what he was doing)		Unclear (did not explain what he was doing)
1	2	3	4	5
Not Understanding (lacked sympathy or empathy; was intolerant; did not reassure)		Fairly Understanding (generally sympathetic or empathic; tolerant; reassuring)		Very Understanding (very sympathetic or empathetic; very tolerant; very reassuring)
1	2	3	4	5
Secure (very relaxed and confident)		Average (reasonably relaxed)		Insecure (puzzled and hesitant)
1	2	3	4	5
Impatient (interrupted me; was in a hurry)		Average (interrupted me occasionally; somewhat hurried)		Patient (did not interrupt me--"heard me out"; was unhurried)
1	2	3	4	5
Very Gentle During the Physical Exam	Reasonably Gentle During the Physical Exam	Careful During the Physical Exam	A Little Rough During the Physical Exam	Very Rough During the Physical Exam

APPENDIX D

TEST OF COGNITIVE KNOWLEDGE

IN NEUROLOGY

APPENDIX D

TEST OF COGNITIVE KNOWLEDGE

IN NEUROLOGY

FOR QUESTIONS 1-10

For each of the following questions select the one appropriate answer from the following list and mark the corresponding number on your answer sheet.

1. Cerebellar system
2. Vestibular system
3. Basal ganglion
4. Dorsal columns of cord
5. Pyramidal system

1. Paucity of eyeblinks and masklike facies

2. Abnormal finger to nose test with difficulty in coordinated control noted near reaching endpoint

3. Positive Romberg eyes closed only

4. Rebound phenomenon on releasing restrained flexed forearm

5. Abnormal cold caloric test

6. Positive Romberg eyes open and only slightly accentuated with eyes closed

7. Slowness in starting movement (bradykinesia)

8. Muscle spasticity

9. Muscle rigidity (cog-wheel type)

10. Loss of vibratory sensation

FOR QUESTIONS 11-20:

Select the most appropriate type of general lesion for each of the following questions from the four listed below and mark the corresponding number on your answer sheet.

1. Lower motor lesion
2. Upper motor lesion
3. Fairly equally applies to both upper and lower motor lesion
4. None of the above

11. Flaccidity of muscles

12. Fasciculations

13. Babinski sign

14. Ankle Clonus

15. Hyperreflexia

16. Rossolimo sign

17. Difficulty in closing the right eye and drooping of the right corner of the mouth

18. Weakness

19. Spasticity of muscles

20. Hyperactive jaw reflex

FOR QUESTIONS 21-25:

For the following items select the one most appropriate answer from the list below in deciding on a diagnosis of primary muscle disease or muscle dysfunction secondary to a peripheral nerve or lower motor neuron lesion.

1. Muscle dysfunction secondary to a peripheral nerve or lower motor neuron lesion.
2. Primary muscle disease
3. Found fairly equally in both 1 and 2
4. None of the above

21. Frequent fibrillations noted on EMG studies

22. Muscle wasting

23. Objective loss of pain and light touch sensations associated with muscle dysfunction

24. Frequent fasciculations observed on gross inspection

25. Retained deep tendon reflexes despite marked loss in muscle bulk and weakness.

FOR QUESTIONS 26-35:

In the following questions select from the list below the one most likely site where a single lesion could produce the described symptomatology. Mark your answer sheet with the appropriate number.

1. Lumbar cord or nerve root
2. Thoracic cord or nerve root
3. Cervical cord or nerve root
5. Brain stem
6. Cerebral cortex or internal capsule

26. Difficulty in abducting the left eye with weakness of the right hand and leg and a right Babinski.

27. Ptosis with outward and downward deviation of the left eye as well as weakness of the right arm and leg and a right Babinski.

28. Inability to close eye on the left and drooping of the left corner of the mouth with a right Babinski.

29. Loss of sensation to pinprick and light touch on the right side of the face and left side of the body with slight ataxia and falling to the right.

30. Extreme pain radiating down left lower extremity.

31. Right hemiparesis including the lower half of the face on the right and a right Babinski.

32. Atrophy of biceps and deltoid on the right and small muscles of both hands with fasciculations of the involved muscles and bilateral ankle clonus and hyperreflexia.

33. Inability to speak with no paralysis of the muscles of speech (expressive aphasia).

34. Inability to differentiate size of coins and texture of materials with the right hand with no loss of light touch.

35. Deviation of the tongue to the right with frequent fasciculations as well as a left Babinski and Hoffman sign.

APPENDIX E

NEUROLOGICAL PSYCHOMOTOR

PRACTICAL TEST

APPENDIX E

NEUROLOGICAL PSYCHOMOTOR

PRACTICAL TEST

Each test will be evaluated in the three following areas:

A. Patient rapport with emphasis on accuracy of instructions given to patient before and during each test.

B. Proper positioning of patient.

c. Proper technique in performing the test.

Each of the above categories will be graded as follows:

3 = Excellent

2 = Adequate

1 = Fail

The three categories are combined for a total score. The total score represents the student's psychomotor skill or technique in performing the various neurological tests.

STUDENT NAME ____________________

PRACTICAL NEUROLOGICAL PSYCHOMOTOR TESTING

	Accuracy of Instructions	Positioning of Patient	Technique	TOTAL
	A	B	C	
PATELLAR (sitting)				
(lying)				
ACHILLES (sitting)				
(lying)				
(kneeling)				
BICEPS				
TRICEPS				
BRACHIORADIALIS				
HOFFMAN				
PLANTAR				
ANKLE CLONUS				
ABDOMINALS				
MUSCLE TONE (upper extremity)				
MUSCLE STRENGTH				
MUSCLE STRENGTH				
VIBRATION				
PAIN (specific area)				
LIGHT TOUCH (specific area)				
POSITION SENSE				
STEREOGNOSIS				
CRANIAL NERVE				
CRANIAL NERVE				
CEREBELLAR FUNCTION (upper extremity)				
(upper extremity)				
(lower extremity)				
(lower extremity)				
TOTAL SCORE				

APPENDIX F

PHYSICAL DIAGNOSIS

PRACTICAL EXAMINATION RATING SHEET

APPENDIX F

PHYSICAL DIAGNOSIS

PRACTICAL EXAMINATION RATING SHEET

(Circle All Ratings Assigned)

Student:__________________ Patient:________________ Physician:_______________

	(5) Superior: Needs No Supervision	(4)	(3) Satisfactory: Needs Supervision	(2)	(1) Inadequate or Omitted
HISTORY:					
1. Chief complaint	5	4	3	2	1
2. Social history	5	4	3	2	1
3. Previous medical history	5	4	3	2	1
4. Family history	5	4	3	2	1
5. Systems review	5	4	3	2	1
PHYSICAL EXAMINATION:					
1. General skills	5	4	3	2	1
2. Head and neck	5	4	3	3	1
3. Thorax	5	4	3	2	1
4. Abdomen	5	4	3	2	1
5. Extremities, skin, and appendages	5	4	3	2	1
6. Neurologic	5	4	3	2	1
7. Structural	5	4	3	2	1

	(5) Superior	(4)	(3) Satisfactory	(2)	(1) Inadequate
DOCTOR-PATIENT RELATIONSHIP: During the interview and examination, was the student:					
1. relaxed with the patient?	5	4	3	2	1
2. interested in his problems?	5	4	3	2	1
3. tactful in discussing sensitive areas?	5	4	3	2	1
4. able to keep the patient at ease?	5	4	3	2	1

APPENDIX G

STUDENT RATING OF SIMULATED PATIENT EXAMINATION EXPERIENCE

APPENDIX G

STUDENT RATING OF SIMULATED PATIENT EXAMINATION EXPERIENCE

Now that you have completed the complete neurological evaluation history and physical examination, please complete the following:

How would you describe the simulated patient examination experience?

Useful	1	2	3	4	5	6	7	Useless
Meaningful	1	2	3	4	5	6	7	Meaningless
Important	1	2	3	4	5	6	7	Unimportant
Realistic	1	2	3	4	5	6	7	Artificial
Successful	1	2	3	4	5	6	7	Unsuccessful

How would you describe your feelings about yourself now that you have performed the simulated patient examination?

Secure	1	2	3	4	5	6	7	Insecure
Successful	1	2	3	4	5	6	7	Unsuccessful
Calm	1	2	3	4	5	6	7	Anxious
Pleasant	1	2	3	4	5	6	7	Unpleasant
Competent	1	2	3	4	5	6	7	Incompetent

APPENDIX H

INTERCORRELATION MATRIX FOR ALL VARIABLES IN CHAPTER IV

APPENDIX H

INTERCORRELATION MATRIX FOR ALL VARIABLES IN CHAPTER IV

Chapter IV reports the analysis of data. Nineteen variables are identified. This appendix presents the intercorrelation matrix for all 19 variables. Below mnemonic terms are listed together with the variables they represent.

1. NEHPE--means T-score for two complete neurological evaluation history and physical examinations
2. PR-NEHPRE--mean T-score for two patient ratings of two complete neurological evaluation history and physical examinations
3. NAMID--neuroanatomy mid-term test
4. NSMID--neurosciences mid-term test
5. NAFIN--neuroanatomy final test
6. NSFIN--neurosciences final test
7. NAPRAC--neuroanatomy practical test
8. TVCAS1--TV case evaluation I
9. TVCAS2--TV case evaluation II
10. PMP--Patient Management Problem
11. CONPRN--measures of concepts and principles in the Neuromuscular Instructional System (3 through 7 above)

12. CLNSKL--measures of clinical skills in the Neuromuscular Instructional System (8 through 10 above)

13. GRADES--overall weighted T-score derived from 3 through 10 above

14. GRADES*--overall weighted T-score derived from 3 through 10 above plus 19 below

15. COG--test of cognitive knowledge in neurology

16. HIST-PD--rating of history-taking skill from a physical diagnosis course

17. PE-PD--rating of physical examination skill from a physical diagnosis course

18. DPR-PD--rating of doctor-patient relationship skill from a physical diagnosis course

19. PSYMO--neurological psychomotor practical test

APPENDIX H

INTERCORRELATION MATRIX FOR ALL VARIABLES IN CHAPTER IV

	1 NEHPE	2 PR-NEHPE	3 NAMID	4 NSMID	5 NAFIN	6 NSFIN	7 NAPRAC	8 TVCAS1	9 TVCAS2	10 PMP	11 CONPRN	12 CLNSKL	13 GRADES	14 GRADES*	15 COG	16 HIST-PD	17 PE-PD	18 DPR-PD	19 PSYMO
1. NEHPE	1.00																		
2. PR-NEHPE	.198	1.00																	
3. NAMID	.341	.250	1.00																
4. NSMID	.310	.248	.519	1.00															
5. NAFIN	.385	.288	.732	.674	1.00														
6. NSFIN	.392	.229	.634	.555	.642	1.00													
7. NAPRAC	.319	.123	.137	.301	.378	.280	1.00												
8. TVCAS1	.185	.107	.254	.439	.421	.364	.174	1.00											
9. TVCAS2	.243	.132	.373	.345	.460	.320	.180	.323	1.00										
10. PMP	.182	.060	.263	.432	.393	.358	.348	.191	.042	1.00									
11. CONPRN	.462	.299	.734	.834	.874	.779	.553	.473	.433	.519	1.00								
12. CLNSKL	.265	.144	.532	.585	.666	.561	.250	.736	.673	.570	.690	1.00							
13. GRADES	.457	.258	.699	.795	.838	.769	.494	.554	.490	.555	.955	.781	1.00						
14. GRADES*	.459	.299	.707	.826	.867	.756	.528	.554	.475	.543	.976	.766	.954	1.00					
15. COG	.172	.119	.251	.286	.342	.311	.081	.215	.099	.114	.326	.238	.378	.330	1.00				
16. HIST-PD	.258	-.093	.142	.082	.044	.114	.116	.051	-.051	-.054	.103	-.011	.090	.058	-.087	1.00			
17. PE-PD	.194	-.025	.100	.030	.025	.075	.043	.003	.021	-.049	.062	-.017	.052	.094	-.111	.006	1.00		
18. DPR-PD	.019	.044	.022	.084	-.105	.037	-.067	.140	.024	-.021	.001	.072	.033	.006	-.007	.274	.050	1.00	
19. PSYMO	.302	.245	.300	.463	.442	.338	.350	.287	.081	.227	.499	.273	.461	.625	.216	-.081	.245	-.043	1.00

LIST OF REFERENCES

LIST OF REFERENCES

Ad Hoc Committee on Evaluation Procedures. Final report of the Ad Hoc Committee on Evaluation Procedures. East Lansing: College of Ostepathic Medicine, Michigan State University, 1975. (Document)

American Association of Colleges for Teacher Education. A survey of colleges of teacher education. In A. A. Schmieder (ed.). *Comptency Based Education: The State of the Scene.* Washington, D.C.: American Association of Colleges for Teacher Education, 1973.

Barrows, H. S., and Abrahamson, S. The programmed patient: A technique for appraising student performance in clinical neurology. *Journal of Medical Education*, 1964, *39*, 802-805.

Bartlett, J. W. Medical school and career performance of medical students with low Medical College Admission Test scores. *Journal of Medical Education*, 1967, *42*, 231-237.

Beatty, J. D., and Cleeton, G. U. Predicting achievement in college and after graduation. *Personnel Journal*, 1928, *6*, 344-351.

Best, W. R., Dekma, A. J., Fisher, L. A., and Smith, N. E. Multivariate predictors in selecting medical students. *Journal of Medical Education*, 1971, *46*, 42-50.

Block, J. H. *Mastery learning: Theory and practice.* New York: Holt, Rinehart, and Winston, 1971.

Bloom, B. S. Learning for mastery. UCLA-CSEIP *Evalaution Comment*, 1968, *1*(2).

Bracht, G. H., and Glass, G. V. The external validity of experiments. *American Educational Research Journal*, 1968, *5*, 437-474.

Brandt, E. M., and Metheny, B. H. Relationships between measures of student and graduate performance. *Nursing Research*, 1968, *17*, 242-246.

Bridgman, D. S. Success in college and business. *Personnel Journal*, 1930, *9*, 1-19.

Buehler, J. A., and Trainer, J. B. Prediction of medical school performance and its relationship to achievement. *Journal of Medical Education*, 1962, *37*, 10-18.

Campbell, D. T., and Stanley, J. C. *Experimental and quasi-experimental designs for research.* Chicago: Rand McNally, 1963.

Carnegie Commission on Higher Education. *Higher education and the nation's health: Policies for medical and dental education.* Hightstown, New Jersey: McGraw-Hill, 1970.

Chambers, J. A. Relating personality and biographical factors to scientific creativity. *Psychological Monographs*, 1964, *78*(584).

Chan, S., Parmeter, J. T., and Ways, P. O. Evaluation of students and programs in a multi-community fundamental clerkship. *Association for Hospital Medical Education Journal*, 1974, *7*, 15-21.

Clute, K. F. *The general practitioner: A study of medical education and practice in Ontario and Nova Scotia.* Toronto: University of Toronto Press, 1963.

Coggeshall, L. T. *Planning for medical progress through education.* Evanston, Illinois: Association of American Medical Colleges, 1965.

Cole, D. L. The prediction of teaching performance. *Journal of Educational Research*, 1961, *54*, 345-348.

Committee on Goals and Priorities of the National Board of Medical Examiners. *Evaluation in the continuum of medical education: Report of the Committee on Goals and Priorities of the National Board of Medical Examiners.* Philadelphia: National Board of Medical Examiners, 1973.

Cornfield, J., and Tukey, J. W. Average values of mean squares in factorials. *Annals of Mathematical Statistics*, 1956, *27*, 907-949.

Cox, S. G. Do educational measures predict vocational success? *Vocational Guidance Quarterly*, 1971, *19*, 271-273.

Dimond, E. G. The docent team-Kansas City General Hospital. *Missouri Medicine*, 1972, *69*, 894-896.

Ebel, R. L. *Essentials of educational measurement.* Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1972.

Elstein, A. S., Sprafka, S., and Shulman, L. S. Analyzing medical inquiry processes. East Lansing: Office of Medical Education Research and Development, Michigan State University, 1973. (Document)

Erdmann, J. B. Separating wheat from chaff: Revision of the MCAT. *Journal of Medical Education*, 1972, *47*, 747-749.

Erdmann, J. B., Mattson, D. E., Hutton, J. G., Jr., and Wallace, W. L. The Medical College Admission Test: Past, present, and future. *Journal of Medical Education*, 1971, *46*, 937-946.

Fredericks, M. A., and Mundy, P. The relationship between social class, average grade in college, Medical College Admission Test scores, and academic achievement of students in medical school. *Journal of Medical Education*, 1967, *42*, 126-133.

Froelich, R. E. A course in medical interviewing. *Journal of Medical Education*, 1969, *44*, 1165-1169.

Gambrill, B. L. *College achievement and vocational efficiency.* New York: Teachers College, Columbia University, 1922.

Ginzberg, J. A. L., and Herme, J. L. *Talent and Performance.* New York: Columbia University Press, 1964.

Githens, J. H., Cooper, W. G., and Helfer, R. E. Experience with pre-testing students entering medical school. *Journal of Medical Education*, 1970, *45*, 255-258.

Glaser, R., and Nitko, A. J. Measurement in learning and instruction. In R. L. Thorndike (ed.). *Educational Measurement.* Washington, D.C.: American Council on Education, 1971.

Goran, M. J., Williamson, J. W., and Gonnella, J. S. The validity of patient management problems. *Journal of Medical Education*, 1973, *48*, 171-177.

Goroll, A. H., Stoeckle, J. D., and Lazare, A. Teaching the clinical interview: An experiment with first-year students. *Journal of Medical Education*, 1974, *49*, 957-962.

Gough, H. G., and Hall, W. B. An attempt to predict graduation from medical school. *Journal of Medical Education*, 1975, *50*, 940-950.

Gough, H. G., Hall, W. B., and Harris, R. E. Admission procedures as forecasters of performance in medical training. *Journal of Medical Education*, 1963, *38*, 983-998.

Haley, H. B., Juan, I. R., and Paiva, R. E. A. MCAT scores in relation to personality measures and biographical variables. *Journal of Medical Education*, 1971, *46*, 947-958.

Haley, J. V. Medical School Admission Test as a predictor of grade factors. *Journal of Medical Education*, 1973, *48*, 98-100.

Haley, J. V., and Lerner, M. J. The characteristics and performance of medical students during preclinical training. *Journal of Medical Education*, 1972, *47*, 446-452.

Harmon, L. R. The development of a criterion of scientific competence. In C. W. Taylor and E. Barron (eds.). *Scientific creativity: Its recognition and development*. New York: John Wiley and Sons, 1963.

Harmon, L. R. Progress and potentiality: Career determiners of high-level personnel. In C. W. Taylor (ed.). *Widening horizons in creativity*. New York: John Wiley and Sons, 1964.

Harrell, T. W. High earning MBA's. *Personnel Psychology*, 1972, *25*, 523-530.

Harrell, T. W., and Harrell, M. S. Relation of second year MBA grades to business earnings. *Personnel Psychology*, 1974, *27*, 487-491.

Harrill, T. S. A comparative study of the efficiency of intellective and nonintellective measures as predictors of job success and job satisfaction. Unpublished doctoral dissertation, Auburn University, 1973.

Harris, M. L., and Stewart, D. M. Application of classical strategies to criterion-referenced test construction. Paper presented to the annual meeting of the American Educational Research Association, 1971.

Hess, J. W. A comparison of methods for evaluating medical student skill in relating to patients. *Journal of Medical Education*, 1969, *44*, 934-938.

Hill, J. K. Assessment of intellectual promise for medical school. *Journal of Medical Education*, 1959, *34*, 959-964.

Hiss, R. G., and Peirce, J. C. A strategy for developing educational objectives in medicine: Problem-solving skills. *Journal of Medical Education*, 1974, *49*, 660-665.

Hoffman, E. L., Wing, C. W., and Lief, H. I. Short and long-term predictions about medical students. *Journal of Medical Education*, 1963, *38*, 852-857.

Holt, R. R., and Luborsky, L. *Personality Patterns of Psychiatrists*. (Volumes 1 and 2.) New York: Basic Books, 1958.

Houston, W. R. (ed.). *Exploring Competency based education*. Berkeley, California: McCutchan Publishing Corporation, 1974.

Houston, W. R., and Howsam, R. B. (eds.). *Competency-based teacher education: Progress, problems, and prospects*. Palo Alto, California: Science Research Associates, 1972.

Howell, M. A., and Vincent, J. W. The Medical College Admission Test as related to achievement tests in medicine and to supervisory evaluations of clinical physicians. *Journal of Medical Education*, 1967, *42*, 1037-1044.

Isaac, S., and Michael, W. B. *Handbook in research and evaluation*. San Diego: R. R. Knapp, 1971.

Jason, H. Evaluation of basic science learning: Implications of and for the 'GAP Report'. *Journal of Medical Education*, 1974, *49*, 1003-1004.

Jason, H., Kagan, N., Werner, A., Elstein, A., and Thomas, J. New approaches to teaching basic interview skills to medical students. Paper presented to the annual meeting of the American Psychiatric Association, 1970.

Jepsen, V. L. Scholastic proficiency and vocational success. *Educational Psychological Measurement*, 1951, *11*, 616-628.

Johnson, D. G. A multi-factor method of evaluating medical school applicants. *Journal of Medical Education*, 1962, *37*, 656-665.

Johnson, D. G., and Hutchins, E. B. *Doctor or dropout? A study of medical student attrition.* Evanston, Illinois: Association of American Medical Colleges, 1966.

Jones, M. L. Analysis of certain aspects of teaching ability. *Journal of Experimental Education*, 1956, *25*, 153-180.

Jones, R. D. The prediction of teaching efficiency from objective measures. *Journal of Experimental Education*, 1946, *15*, 85-99.

Kegel-Flom, P. Predicting supervisor, peer, and self ratings of intern performance. *Journal of Medical Education*, 1975, *50*, 812-815.

Kempthorne, O. The design and analysis of experiments with some reference to educational research. In R. O. Collier, Jr., and S. M. Elam (eds.). *Research design and analysis: Second annual Phi Delta Kappa symposium on educational research.* Bloomington, Indiana: Phi Delta Kappa, 1961.

Kerlinger, F. N., and Pedhazur, E. J. *Multiple regression in behavioral research.* New York: Holt, Rinehart, and Winston, 1973.

Kole, D. M., and Matarazzo, J. D. Intellectual and personality characteristics of two classes of medical students. *Journal of Medical Education*, 1965, *40*, 1130-1144.

Korman, M., and Stubblefield, R. L. Medical school evaluation and internship performance. *Journal of Medical Education*, 1971, *46*, 670-673.

Kunkel, B. W. Standing of undergraduates and alumni. *School and Society*, 1917, *5*, 549-555.

Lins, L. J. The prediction of teaching efficiency. *Journal of Experimental Education*, 1946, *15*, 2-60.

McGuire, C., and Page, G. An overview of physicians' performances on simulated clinical encounters presented in patient management problem format. Paper presented to the annual meeting of the American Educational Research Association, 1973.

McGuire, C., and Solomon, L. *Clinical simulations*. New York: Appleton-Century-Crofts, 1971.

Magnusson, D. *Test theory*. Reading, Massachusetts: Addison-Wesley, 1967.

Mandell, M. Selecting chemists for the federal government. *Personnel Psychology*, 1950, *3*, 53-56.

Martin, R. A., and Pacheres, J. Good scholars not always best. Cited in research section of *Business Week*, February 24, 1962, 77-78.

Massey, H. W., and Vineyard, E. E. Relationship between scholarship and first year teaching success. *Journal of Teacher Education*, 1958, *9*, 297-301.

Medical school admission requirements, U.S.A. and Canada 1975-1976. Washington, D.C.: Association of American Medical Colleges, 1974.

Mehrens, W. A., and Lehmann, I. J. *Measurement and evaluation in education and psychology*. New York: Holt, Rinehart and Winston, 1973.

Miller, G. E. *Teaching and learning in medical school* Cambridge, Massachusetts: Harvard University Press, 1961.

Millis, J. S. *A rational public policy for medical education and its financing*. New York: The National Fund for Medical Education, 1971.

Moffatt, D. J., Jacobs, A. W., and Metcalf, W. K. Predictors of academic performance in gross anatomy. *Journal of Medical Education*, 1971, *46*, 545-548.

Monk, M. A., and Thomas, C. B. Characteristics of male medical students related to their subsequent careers. *Johns Hopkins Medical Journal*, 1970, *127*, 254-272.

Morris, W. W. *Problems in medical student selection.* Evanston, Illinois: Association of American Medical Colleges, 1958.

Muchinsky, P. M., and Hoyt, D. P. Academic grades as a predictor of occupational success among engineering graduates. *Measurement and Evaluation in Guidance*, 1973, *6*, 93-103.

Nair, S. R. Academic achievement and social background as predictors of teacher effectiveness. *Teacher Education*, 1973, *7*, 17-22.

National Board of Examiners for Osteopathic Physicians and Surgeons, Inc. *Journal of the American Osteopathic Association, Education Annual 1974*, *73* (Supplement), 64-68.

Neufeld, V. R., and Barrows, H. S. The 'McMaster philosophy': An approach to medical education. *Journal of Medical Education*, 1974, *49*, 1040-1050.

Pallett, J. E. Definition and prediction of success in the business world. Unpublished doctoral dissertation, University of Iowa, 1965. Cited in D. P. Hoyt, The relationship between college grades and adult achievement: A review of the literature. *ACT Research Reports*, No. 7. Iowa City: American College Testing, 1965.

Pallett, J. E., and Hoyt, D. P. College curriculum and success in general business. *Journal of College Student Personnel*, 1968, *9*, 238-245.

Payne, G. E. Scholarship and success in teaching. *Journal of Educational Psychology*, 1918, *9*, 217-219.

Penta, F. B., and S. Kofman. The effectiveness of simulation devices in teaching selected skills of physical diagnosis. *Journal of Medical Education*, 1973, *48*, 442-445.

Peterson, O. L., Andrews, L. P., Spain, R. S., and Greenberg, B. G. An analytic study of North Carolina general practice: 1953-1954. *Journal of Medical Education*, 1956, *31*, Part 2.

Pierson, G. A. School marks and success in engineering. *Educational and Psychological Measurement*, 1947, *7*, 612-617.

Price, P. B., Taylor, C. W., Richards, J. M., and Jacobsen, T. L. *Measurement and prediction of physician performance.* Salt Lake City: L.L.R. Press, 1971.

Price, P. B., Taylor, C. W., Richards, J. M., and Jacobsen, T. L. Measurement of physician performance. *Journal of Medical Education*, 1964, *39*, 203-211.

Rice, D. E. A study of incomes of technically trained men. *Scientific American*, 1913, *109*, 116-117.

Richards, J. M., Taylor, C. W., and Price, P. B. The prediction of medical intern performance. *Journal of Applied Psychology*, 1962, *46*, 142-146.

Richards, J. M., Taylor, C. W., Price, P. B., and Jacobsen, T. L. An investigation of the criterion problem for one group of medical specialists. *Journal of Applied Psychology*, 1965, *49*, 79-90.

Roemer, R. E. Nine year validity study of predictors of medical school success. *Journal of Educational Research*, 1965, *59*, 183-185.

Rothman, A. I. A comparison of persistent high and low achievers through four years of undergraduate medical training. *Journal of Medical Education*, 1973, *48*, 180-182.

Rothman, A. I., and Flowers, J. F. Personality correlates of first-year medical school achievement. *Journal of Medical Education*, 1970, *45*, 901-905.

Saffer, J. B., and Saffer, L. D. Academic record as a predictor of future job performance of nurses. *Nursing Research*, 1972, *21*, 457-462.

Scheifley, V. M., and Schmidt, W. H. Jeremy D. Finn's multivariance--univariate and multivariate analysis of variance, covariance, and regression (modified and adapted for use on the CDC 6500). *Occasional Paper No. 22.* East Lansing: Office of Research Consultation, College of Education, Michigan State University, 1973.

Schick, G. J. The predictive value of a teacher judgment test. Unpublished doctoral dissertation, University of Wisconsin, 1957. Cited in D. P. Hoyt, The relationship between college grades and adult achievement: A review of the literature. *ACT Research Reports*, No. 7. Iowa City: American College Testing, 1965.

Selover, R. B. Identifying college graduates with high potential for advancement. Unpublished mimeographed report, Prudential Insurance Company, 1962. Cited in J. P. Campbell, M. D. Dunnette, E. E. Lawler, and E. E. Weick. *Managerial behavior, performance, and effectiveness.* New York: McGraw-Hill, 1970.

Simun, P. B., and Asher, J. W. The relationship of variables in undergraduate school and school administrators' ratings of first-year teachers. *Journal of Teacher Education*, 1964, *15*, 293-302.

Stead, E. A., Jr. Conserving costly talents--providing physicians' new assistants. *Journal of the American Medical Association*, 1966, *198*, 182-183.

Stefanu, C., and Farmer, T. A. The differential predictive validity of science MCAT in the admission process. *Journal of Medical Education*, 1971, *46*, 461-463.

Steffenson, J. P. Forward to *Exploring competency-based education.* In W. R. Houston (ed.). *Exploring competency-based education.* Berkeley, California: McCutcheon Publishing Corporation, 1974.

Steiner, B. W., Garfinkel, P. E., and Hunter, R. C. The processes of psychiatric residency training: Part I: Selection of trainees and the outcome of training. *Canadian Psychiatric Association Journal*, 1974, *19*, 193-200.

Stoelting, G. J. The selection of candidates for teacher education at the University of Wisconsin. *Journal of Experimental Education*, 1955, *24*, 115-132.

Stuit, D. B. Scholarship as a factor in teaching success. *School and Society*, 1937, *46*, 382-384.

Talalla, A., Boufford, J. I., and Lass, S. L. An integrated clinical correlation course in the neurosciences for first-year medical students. *Journal of Medical Education*, 1974, *49*, 253-263.

Taylor, C. W. Variables related to creativity and productivity among men in two research laboratories. In C. W. Taylor and F. Barron (eds.). *Scientific creativity: Its recognition and development.* New York: John Wiley and Sons, 1963.

Taylor, C. W., and Ellison, R. L. Predicting creative performance from multiple measures. In C. W. Taylor (ed.). *Widening horizons in creativity.* New York: John Wiley and Sons, 1964.

Taylor, C. W., Price, P. B., Richards, J. M., and Jacobsen, T. L. An investigation of the criterion problem for a medical school faculty. *Journal of Applied Psychology*, 1964, *48*, 294-301.

Taylor, C. W., Price, P. B., Richards, J. M., and Jacobsen, T. L. An investigation of the criterion problem for a group of medical general practitioners. *Journal of Applied Psychology*, 1965, *49*, 399-406.

Taylor, C. W., Smith, W. R., and Ghiselin, B. The creative and other contributions of one sample of research scientists. In C. W. Taylor and F. Barron (eds.). *Scientific creativity: Its recognition and development.* New York: John Wiley and Sons, 1963.

Taylor, C. W., Smith, W. R., Ghiselin, B., and Ellison, R. *Explorations in the measurement and prediction of contributions of one sample of scientists.* Report ASD-TR-61-96, Aeronautical Systems Division, Personnel Laboratory, Lackland Air Force Base, Texas, 1961. Cited in D. P. Hoyt, The relationship between college grades and adult achievement: A review of the literature. *ACT Research Reports*, No. 7. Iowa City: American College Testing, 1965.

Tinning, F. C. An experimental study investigating the effects of real and simulated clinical training on psychomotor, affective, and cognitive variables during real clinical performance of first-year osteopathic medical students. Unpublished doctoral dissertation, Michigan State University, 1973.

Tinning, F. C., and Yinger, R. J. Final report to medical schools studies program of the Merck Foundation on simulation in medical education. East Lansing: College of Osteopathic Medicine, Michigan State University, 1975. (Document)

Tremonti, L. P. Curriculum development. *Journal of Medical Education*, 1974, *49*, 444-446.

Turner, E. V., Helper, M. M., and Kriska, S. D. Predictors of clinical performance. *Journal of Medical Education*, 1974, *49*, 338-342.

Turner, E. V., Helper, M. M., Kriska, S. D., Singer, S. A., and Ruma, S. J. Evaluating clinical skills of students in pediatrics. *Journal of Medical Education*, 1962, *47*, 959-965.

Walters, R. W., and Bray, D. W. Today's search for tomorrow's leaders. *Journal of College Placement*, 1963, *24*, 22-23.

Weinstein, A. G., and Srinivasan, V. Predicting managerial success of MBA graduates. *Journal of Applied Psychology*, 1974, *59*, 207-212.

Westling-Wikstrand, H., Monk, M. A., and Thomas, C. B. Some characteristics related to the career status of women physicians. *Johns Hopkins Medical Journal*, 1970, *127*, 273-286.

Williams, F. J., and Harrell, T. W. Predicting success in business. *Journal of Applied Psychology*, 1964, *48*, 164-167.

Wingard, J. R., and Williamson, J. W. Grades as predictors of physicians' career performance: An evaluative literature review. *Journal of Medical Education*, 1973, *48*, 311-322.