DOCUMENT RESUME

ED 219 009 HE 015 251

AUTHOR Patel, Vimla; Cranton, Patricia A.
TITLE An Alternative Study of Transfer of Learning in Clinical Evaluation.
PUB DATE Mar 82
NOTE 37p.; Paper presented at the Annual Meeting of the American Educational Research Association (New York, NY, March 19-23, 1982).

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Behavioral Objectives; Classroom Observation Techniques; *Clinical Experience; Cognitive Objectives; *Educational Environment; Educational Research; Ethnography; Higher Education; Intellectual Disciplines; Interpersonal Competence; Knowledge Level; Learning Experience; *Medical Education; Performance; Problem Solving; Research Methodology; *Skill Development; Specialization; Student Attitudes; *Transfer of Training

ABSTRACT

The use of an alternative methodology to study transfer of learning in clinical instruction during medical school was investigated. The environment in which clinical instruction takes place was examined, after which hypotheses were proposed and tested in a quasi-experimental design. The first phase of the study, an ethnographic analysis of the teaching and learning environment, revealed that different disciplines (medicine, pediatrics, and surgery) provided quite different environments for the students. Different types of learning (factual knowledge, problem-solving, technical skill, interpersonal skill, and attitude toward health care) were emphasized in each discipline; students were involved in varying activities for varying amounts of time; and the roles of the individuals involved in the teaching process (residents, interns, attending staff) differed among the rotations. Both of the following hypotheses were partially confirmed: (1) the different disciplines in the clerkship program would facilitate different types of learning, and (2) there would be no significant transfer of learning among the disciplines. Factual knowledge and problem-solving were facilitated by different disciplines. These two types of learning were also the ones that appeared to transfer from one discipline to the other. The remaining types (interpersonal skill, technical skill, and attitude toward health care) showed no evidence of being transferred. These results indicate that the clinical teaching and learning process is complex, and it is suggested that educators consider the variety of tasks that students undertake, the situations encountered, and the types of abilities required. A bibliography is appended. (SW)

An Alternative Study of Transfer of Learning in Clinical Evaluation

Vimla Patel, Ph.D.
Centre for Medical Education, McGill University
Montreal, Quebec

Patricia A. Cranton, Ph.D.
Centre for Teaching and Learning Services, McGill University
Montreal, Quebec

Generally, investigations of the teaching and learning processes have been conducted in the traditional classroom setting, and consequently have emphasized the cognitive domain of student learning (recall of facts, analysis and synthesis of concepts, etc.) However, in higher education, particularly in the health professions, a great deal of instruction takes place in "natural" settings (hospitals, social service agencies, and other institutions). Commonly referred to as clinical teaching, this type of instruction must also be concerned with the other two domains of learning: psychomotor (the performance aspect of various technical and interpersonal skills), and affective (attitudes, emotions, values). In the past it has been assumed that student learning in the clinical area is transferred from one setting to another, from one domain to another, and from one discipline to another. Research in other areas of education has shown that different types of learning and different subject areas can be quite independent of each other. The purpose of this study was to investigate the transfer of learning among domains and disciplines in medical education. Due to the complexity of this particular instructional environment it was necessary to conduct the investigation in two phases. First, an ethnographic approach was used to describe the clinical environment in three specific

disciplines. This description resulted in the generation of hypotheses regarding the transfer of student learning among domains and disciplines. The second phase of the study involved the development of data gathering techniques and the use of a quasi-experimental design to test the generated hypotheses.

Research on Clinical Training Programs

A limited amount of research has investigated the nature of clinical training and the reported results have tended to be inconsistent.

Levine and his colleagues (1975) at the University of Texas Medical School found that in Pediatrics a substantial portion of the cognitive information required for success in the clinical years was acquired prior to entry to the clerkship program. Scores of students entering their last clerkship were almost the same on the pretest in Pediatrics as those of students who took the pretest ten months earlier. Post-test scores were higher than pretest scores for all groups of students. This indicates that although students gained cognitive knowledge in Pediatrics, there was no transfer of cognitive knowledge from other disciplines to Pediatrics. In this case, computerized multiple choice questions were used to test knowledge.

In a study on the effect of previous clerkship training experience on student performance in Internal Medicine at the University of Ottawa Medical School (Smiley, 1979), neither the scores on ward performance nor on multiple choice questions revealed differences attributable to previous clerkship experience. Both tests were administered at the end of the program; however, no pretests were used. The authors concluded that despite areas of instruction and experience which overlap

between Medicine and other disciplines, previous clerkship experience does not have a beneficial effect on student performance in Internal Medicine.

The analysis of posttest data alone is not altogether statistically appropriate (Jennings, 1979). Morse and his co-workers (1975) analyzed the results of a medical pre-clerkship test given to four groups of students before their basic medical clerkship; group one had no previous clerkship, while groups two, three and four completed three, six and nine weeks respectively, of clerkship in a discipline other than Internal Medicine. No differences could be shown between the scores of the four groups.

In contrast to these findings, Schwartz and his co-workers (1974) found that groups of students demonstrated increased surgical knowledge at the time of entering their core surgical clerkship which was influenced by prior exposure to other clinical disciplines, especially Internal Medicine. Further conflicting results were obtained by an investigator at McGill University (Scott, 1970), who used multiple choice questions and patient management problems in the Internal Medicine clerkship to show a positive transfer of knowledge and problem solving ability from other disciplines to Internal Medicine.

These studies raise some methodological issues: the importance of using pretests, and the need to assess the reliability and validity of instruments. There is also a need to look more closely at each of the components of student learning. Are some components of learning more appropriately taught in specific disciplines? Are knowledge, skills and attitudes discipline-specific? The major limitations of the research to date are that clinical evaluation is viewed as a single concept and clinical learning is viewed as one type of learning.

Researchers are beginning to realize that there can be no one strategy for assessing student learning and that the instructor-student- environment interaction is a complex one demanding a comprehensive analysis.

Review of Research Methodology

There has been concern expressed recently regarding the ability of psychology to deal with significant behavioural issues at an adequate level of complexity. The adequacy of traditional experimental research procedures for yielding generalizations beyond the particular research paradigm has been questioned (Argyris, 1975; Bronfenbrenner, 1977; Cronbach, 1975; Ellsworth, 1977). Psychology has tended to utilize a paradigm in which most variables are held constant or are assumed to be constant, and by systematic variation of the variables of interest, the consequent changes in those variables can be observed. The experimenter may vary two or more independent variables simultaneously to investigate not only the independent main effects but also interactions among them. However, these methods may be inadequate for the investigation of complex processes.

From time to time, attention has been directed toward the relative usefulness of laboratory studies and field studies. Brunswik (1956) points out that a systematic design almost invariably involves the use of atypical backgrounds for the behaviours being investigated. This makes it difficult to detect the presence of interacting influences or to understand the magnitude of such influences, since many variables of potential importance have been removed experimentally. By systematically separating the variables affecting behaviour in a natural setting, one is destroying the pattern of relationships among variables as

they exist. If one wishes to generalize the results of behavioural research to the contexts in which the behaviour normally occurs, the pattern of relationships that exists in such contexts must not be neglected.

The comparative experimental approach has been traditionally used to study thinking processes in different cultures, where skills and knowledge are compared between two cultures with little attention paid to the situation in which they occur. However, the variance observed in thinking processes among cultures can be interpreted as differing applications of universal cognitive skills in specific contexts rather than reflecting differing cognitive capabilities, i.e. skills are available but the contexts do not necessarily trigger their use. It becomes important then, to conduct <u>ethnographic</u> analyses prior to experimentation in order to identify the kinds of activities in which people often engage. Neither the ethnographical nor the experimental approach alone is sufficient. Ethnography provides powerful suggestions about the form in which experiments should be run, and some of the major contextual variables that should guide the design of the experimental investigations.

Variations in common activities can be investigated within a given "cultural" setting and on this basis, statements can be made about variations in abilities. This enterprise is basically ethnographic. The analysis could be extended to cross-cultural comparisons by examining different cultures. The general hypothesis is that different cultures provide for different learning experiences and the tasks that a culture frequently poses for its members will be the ones with which they deal effectively. This hypothesis implies that any behaviour is influenced by the situation in which it occurs (Brunswik, 1956; Hammond, 1966).

On both methodological and theoretical grounds, then, the goal is to describe the aspects of the environment which could be expected to influence the way in which individuals engage in such activities as problem solving, rule learning, and learning and remembering skills. These activities can then be studied in specific experimental situations. Therefore, at the very least, theories concerning any general process must contain statements that indicate when the process will be brought to bear on the problem and when it will not. This is the issue raised when it is said that cultural differences in cognition may more nearly reflect changes in the situations to which various cognitive skills are applied, than they do general processes.

To assess the clinical performance of medical students, it is suggested that the different disciplines such as Surgery, Psychiatry, and Medicine be viewed as "culturally" whole. Thus one would need to conduct observational studies within each of these disciplines to establish the types of activities and abilities which are required within each of the "cultural contexts." On the basis of these observational studies, a model could be constructed that describes the types of tasks in particular situations that students should be able to perform competently in different disciplines. With such a model, performance can be conceptualized as the interaction between the natural tasks and the clinical situations. The tasks include the functions that a physician must perform in providing care to a patient e.g., history taking, physical examination, use of laboratory tests, defining clinical problems and management, patient education, etc. The abilities and behaviours needed to perform these tasks include knowledge/understanding, problem solving/clinical judgement, technical skill, interpersonal skill

and a sound attitude toward patients and health care. Descriptions of <u>clinical situations</u> include characteristics of the patients, who may be chronically ill, terminally ill, or recently well patients, and information about the settings in which these patients are seen; e.g., outpatient, inpatient, emergency patient, etc. Figure 1 depicts a model consisting of tasks, abilities and situations. (cf. Burg and Lloyd, 1979).

In measuring the performance of a clinical clerk, the model could be applied to different disciplines, using as many components of competence as appropriate and feasible. Selection of components can be made on the basis of frequency of occurence and importance, as judged by subject area experts. The attention to situational factors in this research approach allows the investigator to ask questions which are more relevant to educational practice. On the other hand, the traditional research approach poses questions that force an isolation of variables such as: What is the optimal clerkship training for medical students? The proposed alternative paradigm would yield the question: What are the characteristics of clerkship training programs in various disciplines? However, the question will not be answered completely until the researcher uses this information to design and conduct an experimental study. This strategy has been supported in a recent review by Engel and Filling (1981).

In conclusion, a move must be made toward representative sampling from both the population and the environment in order to investigate the nature of instructional programs. There should be a combination of observational (ethnographic) and experimental studies in order to understand how, and how much learning occurs in a complex clinical setting, such as in the health professions. This involves the unification of both quantitative and qualitative

research.

The Assessment of Clinical Performance

Earlier studies on the performance of physicians relied heavily upon the "critical incident" technique, in which hundreds of statements of effective and ineffective physician behaviours were collected and analyzed in order to make a list of the required skills and abilities (Flanagan, 1950; Hubbard, et al., 1965; Sanazaro & Williamson, 1968).

The issues involved in the assessment of clinical performance are the definition of the components of performance and the separation of the measurable components since there is a great deal of overlap between different components. Traditionally, clinical competence has been evaluated by assessing clinicians' cognitive knowledge, using multiple choice questions and a clinical examination, (Bashook, 1976). While deficiencies in the conventional or traditional clinical examination have been identified (Wilson et al., 1979) no attempts have been made to improve the assessment of a student's clinical skills. Medical educators have become increasingly concerned about the methods currently in use for the evaluation of clinical competence.

Clinical competence has also been assessed, solely on the basis of clinicians' problem solving ability without taking into consideration the motor skills or the attitudes (Bashook, 1976). Recent writers have noted that senior medical students, house officers, and even practising doctors are probably often remiss in the techniques of interviewing and examination skills, but are only rarely monitored in these activities (Engel, 1976; McGuire & Rutler, 1976).

In nearly every report, the implicit assumption is that

clinical performance as measured in any of the above learning domains can be generalized across all domains. However, the studies conducted in one learning domain, specific discipline or content area cannot be generalized to the overall clinical area.

A model has been developed for problem-based criterion referenced testing of clinical competence (Newble, 1978) in Australia. Both the content of the examination and the selection of test methods are based on patient problem blueprints, which identify key areas that require testing. The method has potential for assessing competence at the undergraduate and post-graduate levels, however, some limitations have been observed.

Another, well-publicized method designed to assess clinical competence at the bedside comes from England (Harden & Gleeson, 1979). This objective structured clinical examination (OSCE) is broken down into various components , e.g., taking a history, auscultation of the heart, interpretation of an electrocardiogram or coming to a conclusion on the basis of the findings. There are many advantages of the OSCE system: it is more valid than the traditional approach to clinical examinations, and more reliable since the patient and examiner variables are at a minimum. However, the student's knowledge and skills are tested in isolated compartments and the student is not tested on the ability to look at the patient as a whole.

Although there is a continuing effort to define medical competence and to assess that competence adequately by developing measurement techniques, it is clear that no single method is appropriate for assessing all the competencies involved in clinical performance.

Assuming that clinical competence is complex, perhaps the appropriate approach is to measure competency in each of the

domains in which a clinician is required to function. Thus cognitive, psychomotor and affective domains should all be included, and at the undergraduate level a variety of disciplines should be represented.

Conceptual Framework for Assessing Clinical Performance

To illustrate what is meant by competence in medicine and to set the stage for the investigation to follow, a conceptual framework for assessing clinical performance is presented. The framework is adapted from a model developed by the National Board of Medical Examiners as a basis for its Comprehensive Qualifying Evaluation program (Burg & Lloyd, 1979). Competence is described, under this model, in terms of both tasks and clinical situations (or subject matter) (Burg et al.,1976). Tasks are defined as functions that a physician must perform in providing care to a patient including:

1) History taking: eliciting a medical history and interpreting historical data from medical records.
2) Physical examination: conducting various aspects of the physical examination and interpreting the findings.
3) The use of laboratory tests and other investigative techniques: selecting and interpreting tests, both laboratory and psychological.
4) Defining clinical problems: synthesizing information from various data sources including the history, physical and laboratory findings to generate and evaluate diagnostic hypotheses.
5) Management: (a) Non-surgical - selecting and applying a range of preventive and therapeutic interventions and monitoring the patient's progress; (b) Surgical - understanding the

indications of various forms of surgical intervention and monitoring the patient's post-operative progress, (c) Psychological - employing psychotherapeutic interventions, including assisting patients in handling grief, and helping them to cope with common life crises; d) Patient education - providing patients with the information which will help them to cope with illness and comply with management plans.

In performing these tasks, five basic behaviours are involved:

1) Knowledge/understanding: to recall information, to identify the meaning of specific signs and symptoms and the results of clinical investigation, and to understand certain concepts and principles.

2) Problem-solving/clinical judgement: to synthesize information from a variety of sources and generate and evaluate diagnostic hypotheses based on the data.

3) Technical skill: to perform a variety of procedures in which psychomotor skills play an especially important role; for example, the performance of various aspects of a physical examination and conducting a diagnostic procedure such as a lumbar puncture.

4) Interpersonal skill: to listen attentively to and understand the patient's verbal and non-verbal behaviour; to effectively employ verbal and non-verbal behaviour; to use reassurance; to display appropriate behaviour with various patient affects.

5) Attitudes: to demonstrate during the conduct of various professional activities, sensitivity, empathy, concern, objectivity, self-confidence, attention to the continuing needs of patients and their families; to handle professional responsibilities in a manner which will maximize a likelihood

of achieving favourable health care outcomes; to perform in an ethical manner.

The second aspect of the model describing competence is the clinical situation in which the tasks are performed. These situations may involve, for example, patients who are chronic cases or emergency cases, in-patients or out-patients, terminally ill or recently well patients.

Although the ultimate goal might be to delineate every component of competence, to do so would take considerable time and effort to accomplish. A more attainable, immediate goal might be to start with a sample of salient components. A reasonable approach is to consider sampling important diseases or health maintenance activities from the domain of problems of the ill, or from health maintenance activities for a particular discipline. To describe the components of competence, it is necessary to integrate the performance of each of the tasks previously described with the specific clinical situations, as illustrated by the model presented earlier in Figure 1. For each sub-component, an objective can be identified, and an instrument can be designed for its evaluation. For each discipline, the matrix could be completed with as many components of competence as appropriate and feasible.

Some previous research has attempted to utilize this descriptive or ethnographic approach to the study of medical education. Becker and his collegues (1961) differentiated between the characteristics of pre-clinical and clinical study, and outlined the students' work in each clinical speciality, but their discussion of student experiences and attitudes remain at a general level, describing features common to most or all of the training. Similarly, Merton and his colleagues (1957), placed

greater stress on such features as speciality choice, with little discussion of the areas of differentiation within the medical school. Bloom's (1971) description of the State University of New York Medical School also tended to examine characteristics across specialities within the school as a whole, rather than locating shifts in "atmosphere" within it. A study conducted at Edinburgh Medical School by Atkinson (1973), highlighted the differences in learning environments within the medical school. Students' perception of their clinical teaching in Medicine and Surgery were compared by means of a questionnaire. Medical and Surgical rotations differed in a number of ways, including students' self-perception and their reported relations with clinicians and patients. The organization of clinical teaching was also found to differ between Medicine and Surgery.

In summary, the assessment of clinical competence in medical education has tended to treat a clinical program as a homogeneous unit. The research findings to date suggest that more careful note should be taken of differences within the medical school. Therefore, the present study begins with an analysis of the environment in three disciplines: Medicine, Pediatrics and Surgery. This part of the study generated research hypotheses which were then tested in the second phase of the study.

Phase I: Ethnographic Analysis

The ethnographic phase of the study was conducted using observational techniques, student questionnaires and student diaries to examine the learning environments in the Medicine, Surgery and Pediatrics rotations.

A number of days were spent in each rotation, collecting qualitative observational data. A student questionnaire was

developed to assess the contribution of instructors to different types of learning, and the organization of student time during the clerkship rotation. The questionnaire was administered at the beginning of the year in all three disciplines when students entered the clerkship rotations with no prior experience in other clerkship disciplines, and again at the end of the year when students had completed most of the rotations.

Student diaries were kept by eight randomly selected students in each of the three disciplines, giving a total sample of twenty-four. The diary consisted of a record of student activities during the clerkship program, including:

1) Time: amount of time devoted to each activity.
2) Activities: lectures, seminars, history taking, outpatient work, etc.
3) Setting: lecture room, library, operating room, etc.
4) Format: alone, with resident, small group, etc.
5) Rating: perceived educational value of each activity on a five point scale.

Students recorded their daily activities for six days in the middle of their eight week program to minimize the "start-up" and "wind-down" effects associated with the beginning and the end of the rotation. Student diaries were analyzed in terms of the following categories:

1) <u>Patient-care activities</u>:

   a) Initial history and physical
   b) Write up
   c) Chart work (notes, reviews, etc.)
   d) Interaction with hospitalized in-patient
   e) Lab work

2) Joint patient care and educational activities:

   a) Work rounds

   b) Teaching rounds

   c) Grand rounds

3) Educational activities:

   a) Teaching conference

   b) Discussion of medical topic (NOT with patient)

   c) Seminar/lecture

   d) Individual study

4) Mechanical activity: ("scutt work")

5) Leisure time

Results of Ethnographic Analysis

Results from the student questionnaire provided a description of the contribution of the clinical instructors to student learning in each of the disciplines (Figure 2) and a description of the activities of the clinical clerks in each discipline (Table 1). Differences were found between Medicine and Surgery in the student ratings of the contribution of both residents and attending staff, but not interns. Medicine and Pediatrics differed on the ratings of the contribution of interns and attending staff to student learning. Overall, residents were seen to have had the largest role in student learning. As can be seen in Table 1, students indicated that they had more opportunities to participate in activities in Pediatrics and Medicine than in Surgery. The students felt they had learned to handle emotional problems more in Pediatrics than in Medicine and Surgery, including following up patients or seeing families of patients. More time was spent in didactic teaching and case presentations at

conferences in Pediatrics and Medicine respectively. However, students learned to function more as a team in Surgery than in the other two rotations.

Figure 3 shows the students' perceptions of their learning during the clerkship program. In Surgery, students felt that they had learned clinical skill (technical and problem solving) to the greatest extent, followed by factual knowledge and interpersonal skills, and least of all, attitudes. In Pediatrics, acquisition of factual knowledge was rated the highest, followed by interpersonal skills. Clinical skill was rated the lowest in Pediatrics. Students perceived the Medical rotation to be of greatest value in terms of acquisition of clinical skill, factual knowledge, interpersonal skills and attitudes towards health care.

The analyses of student diaries showed that in Pediatrics, students spend 30% of their time in pure educational activities (Table 2). Little time is devoted to such activities in Medicine or Surgery. Most activities related to patient care involving problem solving, history taking and physical examination were most frequent in Medicine followed by Surgery and Pediatrics. Joint educational and patient care activities and mechanical or technical activities were more frequent in Surgery.

Observations of the activities in the rotations confirmed the findings from the diary and the questionnaire. It was found that less time is spent on patients' physical examinations in Surgery than in Medicine and Pediatrics. Diagnosis in Surgery is made with the aid of visual materials such as x-rays, slides and other materials. An important environmental difference between Medicine and Surgery concerns the time spent in the wards by the patients. On the average, surgical patients spend considerably less time in bed than do medical patients and the medical wards usually have

some long-term patients. The pattern of admission and treatment affects the scope of clinical teaching. It is generally the case that by the time they are seen in the teaching rounds, most patients, both in medical and surgical wards, will have received treatment and many of their acute signs and symptoms will have abated.

On the basis of this analysis it appears that Surgery and Medicine provide extremely different contexts, with Pediatrics being more similar to Medicine than Surgery. Since learning involves an interaction between the student and the environment, it is expected that different types of learning will be facilitated in the three disciplines. In the Pediatrics clerkship rotation, students spend more time in didactic lectures and also have more free time to pursue independent study. Students in this discipline interact with children, and thus become aware of their interpersonal skills. In Medicine and Surgery, it is unlikely that interpersonal skills are acquired during the training.

By the time the clinical clerks enter the clerkship program, they have had some exposure to medical and surgical clinical training in their previous year; this is not true for Pediatrics. Since problem solving ability is shown to be related to the number of years of clinical experience (Norman et al., 1980), the acquisition of these skills will likely be less in Pediatrics than in Medicine and Surgery.

Most of the activities in the medical discipline are organized around patient care and little emphasis is given to pure educational activities. Students completing this rotation would be expected to gain more skill in the patient-related area, such as problem solving and technical skill.

The surgical discipline emphasizes technical skills more than

interpersonal skills and if the environment selects and trains specific skills, then this rotation will favour the learning of technical skills.

Because of the overlap of material taught between Medicine, Surgery and Pediatrics, it would be expected that there would be some transfer of knowledge from one discipline to another, particularly between Medicine and Pediatrics, however, this is not likely to be significant statistically.

Hypotheses

The environmental studies in three core disciplines of the Senior Clerkship Program indicated that the learning environment in the Medicine, Pediatrics and Surgery rotations are distinctly different. On this basis, it was hypothesized that:

1) The different disciplines in the clerkship program would facilitate different types of learning, and
2) There would be no significant transfer of learning among disciplines.

Phase II: Experimental Study

This part of the study was conducted in two parts: 1) the development and validation of instruments for performance assessment, and 2) the assessment of performance of senior medical students in various disciplines, in order to test the hypotheses generated in Phase I.

Sample

One hundred and sixty senior medical students in their last year of clinical training participated in this study over a twelve month period. Students were randomly assigned to three teaching hospitals. Three major core disciplines, Surgery, Medicine and Pediatrics were included. There were approximately thirty

students in one clinical rotation in each discipline. Two clinical rotations in each discipline were studied.

Measurement Parameters

Student learning was measured in three learning domains, cognitive, psychomotor and affective. The major parameters were further divided into the categories of factual knowledge, problem solving, technical skill, interpersonal skill and attitude. These categories reflect the abilities a physician is required to have in order to perform the following tasks:

1) recall knowledge,
2) define problems,
3) conduct a physical examination,
4) take a patient's history,
5) manage a patient's health in terms of preventive and continuing care

Most abilities are required for most of the tasks performed by a physician; however, one cannot measure all abilities in performing all tasks. In this study, abilities and tasks were matched such that one particular natural task performed by a physician was chosen to study one particular underlying ability (Table 3).

In selecting type of instruments to be used, two criteria were considered: 1) test specificity: a precise match was made of a single testing method to each task category and 2) practicability: test administration should occur in the existing structure of the clinical setting without undue dislocation to the faculty, students or the patients.

A committee consisting of senior clinical instructors in each of the three disciplines assisted in the development of the instruments. The summary of the precise match between abilities,

tasks and measurements instrument is given in Table 4.

Factual knowledge: A set of thirty to sixty domain-referenced multiple choice questions were developed in each of the three disciplines. The procedure used in the discipline of Medicine illustrates the item development. A specialist in each of the areas of Respiratory, Cardiology, Endocrinology, Central Nervous System, Haematology, Nephrology, Gastroentology, Dermatology, Pharmacology, Rheumatology and Immunology submitted a series of questions which related to the most frequent types of patient problems in the areas. A few not so frequent, but important patient-problem related questions were also included.

Because of the development procedure used, which was directly related to the objectives of the rotation, the instruments can be assumed to have content validity; that is, the questions were developed by the senior clinical teachers with respect to a specific content domain. Test-retest reliablility was calculated by administering the questions to a group of senior medical students a few days apart.

Problem solving: Among the skills of a competent physician are the abilities of quickly surveying information available about a patient and making appropriate decisions based on this information. A typical situation in which these skills are exercised is the review of a patient's chart and this is an important and frequent activity beginning in the clerkship program and continuing through a physician's professional life (Goetz, 1979).

Two to four simulated charts containing patient problems commonly encountered in the specific disciplines were constructed in each discipline. First, a brief statement of the hypothetical patient's problem, the salient clinical aspects in the data, and

the correct solution were developed. Appropriate medical records of real patients were then selected and used as the basis for the simulated charts. "Key" charts identifying the relevant items on the charts were prepared by the senior consultants.

In scoring the charts, the students' judgements were objectively compared with those of the experts by counting errors of omission (where items considered pertinent by the experts were missed by the students), and errors of commission (where students selected items not considered pertinent by the experts). A proficiency score (the percentage of pertinent items identified by the student) was computed.

The charts have content validity since they were based on the charts of real patients. Construct validity was estimated by administering the charts to groups of students, interns, and residents. If the chart review taps skills used in the medical practice, then individuals with more professional experience than students should perform better and since experience is related to pattern recognition, experienced individuals should be able to perform just as well, if not better, in a shorter period of time (Norman et al., 1978).

<u>Technical Skill</u>: Technical skill is referred to as the manual skill of conducting a physical examination. Appropriate checklists of necessary skills which each student should have on completion of a particular discipline in the clerkship program were developed. Space was available at the end of each section for any skill not listed which was specific to a patient. The items on the checklist were divided into the individual content areas of physical examination. Furthermore, each item on the checklist covered a single facet of student behaviour and was described in terms such that an observation of any behaviour could

be reduced to a three-category response, "he did it", "he did it well" and "he didn't do it". The total score was the number of items on the checklist which the examiner felt appropriate for a particular patient.

In order to determine the inter-rater reliability of the instruments, eight clinicians scored a videotape of a student performing a patient examination. Test-retest reliability was estimated by asking the raters to assess the same examination at two different times.

In Surgery, it is not a very common practice for a clinical clerk to do a physical examination; most diagnoses are based on slides and x-ray materials. It was necessary to use an appropriate technique to evaluate the technical skill in Surgery, therefore, a series of oral examinations based on slides and x-rays were designed. Different sections of the discipline were examined by experts in the particular areas. The questions were pre-designed and a group of senior students not in the experimental group were scored objectively by use of a grid indicating correct and incorrect answers and scored subjectively by the content experts. This was done to calculate inter-rater reliability and to obtain subjective as well as objective assessment.

Interpersonal Skill: The instrument for measuring interpersonal skill was adapted from the Johns Hopkins Interpersonal Skills Assessment method which was developed to assess students' observation and interpretation of a variety of discrete interpersonal behaviours during history taking (Grayson, 1971). The Hopkins instrument was modified to suit the objectives of the clerkship program under investigation. The multiple choice questions following the interviews were modified accordingly.

The instrument included four videotaped interview segments between health practitioners and patients with trained actors playing both roles. The health practitioners in each scene portrayed appropriate and inappropriate interpersonal behaviours. These behaviours were drawn from five categories:

1) social amenities,
2) sensitivity to patient's feelings,
3) interchange of information,
4) organization and structure of the interview, and
5) attention to environmental factors.

In each scene the patient presented a health problem common in primary care medicine, e.g., back pain, localized swelling, venereal disease, headache. The patients simulated affective states frequently observed in patients namely, pain, depression, fear and anxiety. Following each interview segment, students were asked multiple choice questions, requiring them to identify appropriate and inappropriate behaviours and to explain the consequences of the behavioural techniques shown in the interviews. In addition, students were asked to rate each interview segment on a scale of one to ten to elicit their evaluation of the overall quality of each interaction.

The content validity of the interview segments and the related question were checked by a team of psychologists, general medical practitioners, sociologists and clinical instructors in terms of their relevance to a health practitioner. Test-retest reliability of the instrument was measured by administering the tests to a group of senior medical students a few days apart. There is no evidence to show that when students recognize "good" or "bad" behaviours in a simulated situation they would apply this knowledge in a natural setting. Thus, to validate student

performance on the simulated cases, interpersonal behaviours which were recorded during regular ward assessment of history taking were used. Patients were also asked to rate the students on their performance.

Attitude Towards Health Care: A questionnaire on student attitude towards health care (professional and ethical) was developed. A similar questionnaire was used by the University of Colorado in their Comprehensive Health Care Program (Hammond & Kern,1959). Although the style of the questionnaire was adapted from the above study, the content was developed according to the criteria for health care in the clerkship training program. The subscales were: compassion, patient's rights, geriatric care, psychiatric care, assuming responsibility, preventive medicine and working as a team. To validate the questionnaire, a subjective rating of student attitude during the ward assessment was used. The questionnaire was circulated to all clinical instructors for their assessment of its relevance to the clerkship training. Inter-item correlations were calculated for the items within the subscales of the questionnaire.

Design of the Study

In order to test the generated hypotheses the following procedure was followed.

Cognitive Domain: Multiple choice questions were administered to the students at the beginning and at the end of the discipline. To assess problem solving, the two charts were administered in each of the three disciplines at the beginning and at the end of the discipline. Students were further divided into two groups, each group completing the charts in a different order, to control for variation due to the sequencing of charts. A time limit was set for the individual charts.

Psychomotor Domain: In order to assess students' technical skills, the clinical teachers in each of the two disciplines (Medicine and Pediatrics) selected a patient from the ward and checked relevant items of physical examination for the particular patient. The student, accompanied by the instructor, examined the patient and was scored by the instructor. A subjective written evaluation was given together with the objective checklist.

Affective Domain: Interpersonal skills were measured using the videotape of simulated doctor-patient interviews. Students were divided into two groups, each group viewing the interviews in a different order to control for any variation due to sequence of interviews. The attitude questionnaire was completed by the students only at the end of each clinical rotation.

All the tests described except the attitude questionnaire were administered twice in each of the three disciplines: at the beginning of the clerkship program, when the student had no prior experience and then again at the end of the program, when the student had about one year's experience in various disciplines.

Results of Instrument Development

Cognitive Domain: The multiple choice tests of factual knowledge demonstrated test-retest reliability of .92, .89 and .96 for Pediatrics, Medicine and Surgery respectively. The test-retest reliability for the chart reviews were .95, .97 and .62, (Chart 1) and .90, .82, and .70 (Chart 2) for Medicine, Pediatrics and Surgery, respectively. To assess validity, problem solving charts were administered to two groups of residents and interns in Medicine. One group had a set time limit to complete the charts and the other group had no time limit. The results are shown in Tables 5a and 5b. The mean proficiency scores of residents were higher than those of interns which in turn were

higher than the students' scores. In Pediatrics and Surgery, the charts were administered to a group of interns and residents, with set time limits to complete the charts. Both charts in Pediatrics and Surgery demonstrated construct validity (Tables 5c and 5d).

Psychomotor Domain: To examine the inter-rater reliability of the technical skill checklists, a senior clinical clerk was videotaped examining a patient. The clerk was given a half-page introduction to the patient. Three residents and three consultants were asked to view and score the student using the checklist. The inter-rater reliability of the checklist was found to be .89 between the residents; .80 between the consultants, and .67 between the residents and consultants. The residents and consultants were also asked to view the videotape again after one month. The test-retest reliability was found to be .92 for residents and .95 for consultants.

In order to assess the inter-rater reliability of the Surgery orals, a Surgery oral examination was rated subjectively by a senior consulting surgeon, and objectively on the basis of the number of correct and incorrect answers given by the students. On two occasions the inter-rater reliabilities were found to be .79 and .92.

Test-retest reliability on the interpersonal skill instrument was examined by having a group of ten students view the videotape of doctor-patient interviews a few days apart. The test-retest reliabilities for the four cases were found to be between .93 and .97.

Test-retest reliability for the attitude questionnaire was found to be .82. The inter-item correlations (internal consistency) on the six different attitude categories were found to range from .54 to .70.

Results of the Hypotheses Testing

In order to test hypotheses one and two, analyses of covariance were performed separately for each type of learning.

Hypothesis 1: It was predicted that the three disciplines under investigation would facilitate different types of learning. Three two-way analyses of covariance were performed, with the dependent variables being posttest scores on factual knowledge, interpersonal skill and problem solving, and the covariates being the pretest scores for each type of learning. Medicine, Surgery and pediatrics were three levels of the first factor, and the two rotations were two levels of the second factor in each analysis. When a significant difference was found, Duncan's Multiple Range Test was used to test which of the comparisons were contributing to the difference.

Table 6 presents the results of the analysis for factual knowledge.

TABLE 6

ANALYSIS OF COVARIANCE: FACTUAL KNOWLEDGE

Source	DF	Sum of Squares	Mean Square	F	Significance
Model	6	12721.4	2120.2	140.9	.0001
Error	123	1850.0	15.1		

An overall significant difference was found ($p<.0001$), therefore a Duncan's Multiple Range Test was performed. This analysis revealed a significant difference in level of factual knowledge between Medicine and Surgery ($p<.05$) and between Surgery and Pediatrics ($p<.05$). No significant difference was found between rotations. In the area of factual knowledge, Hypothesis 1 was accepted; the disciplines varied in the degree to which they facilitated this type of learning.

Table 7 presents the results of the analysis of covariance

with posttest scores on interpersonal skill as the dependent variable.

TABLE 7

ANALYSIS OF COVARIANCE: INTERPERSONAL SKILL

Source	DF	Sum of Squares	Mean Square	F	Significance
Model	6	7667.0	1277.8	153.3	.0001
Error	95	792.8	8.3		

An overall significant difference was found. However, the Duncan's Multiple Range Test revealed that no difference existed among disciplines, and that the difference was due to the gain in interpersonal skill from the first rotation to the second rotation ($p<.05$). For this type of learning, therefore, Hypothesis 1 was not accepted; there was no variation among disciplines in the degree to which they faciliated student performance.

Results of the same analysis for problem solving are presented in Table 8.

TABLE 8

ANALYSIS OF COVARIANCE: PROBLEM SOLVING

Source	DF	Sum of Squares	Mean Square	F	Significance
Model	6	57168.0	9528.0	145.07	.0001
Error	107	7027.8	65.7		

An overall significant difference was found ($p<.0001$). The Duncan's Multiple Range Test showed that all three disciplines were significantly different from each other ($p<.05$), and that there was no difference between the first and second rotations. Hypothesis 1 is accepted in the area of problem solving; each of the disciplines facilitated learning in this area to a different degree.

In summary, it is generally the case that different types of learning were facilitated by the different disciplines. Gains in

factual knowledge were significantly lower, and gains in problem solving skill were significantly higher in Medicine than in the other disciplines. In the problem solving area, gains were also significantly higher in Surgery than in Pediatrics. The significant difference between the first and second rotation scores in interpersonal skill, with no difference among disciplines, suggests that this type of learning was facilitated by rotations through disciplines other than those under investigation, the most likely possibility being the psychiatry rotation.

Hypothesis 2: It was predicted that there would be no significant transfer of learning among disciplines. Five two-way analyses of covariance were performed with the second session (rotation) pretests being the dependent variable and first session (rotation) pretests being the covariate. If no transfer of learning was occuring, it would be expected that there would be significant differences among disciplines. The results of the analysis with factual knowledge as the dependent variable are presented in Table 9.

TABLE 9

ANALYSIS OF COVARIANCE ON SESSION 2 PRETEST

FACTUAL KNOWLEDGE

Source	DF	Sum of Squares	Mean Square	F	Significance
Model	5	8402.7	1680.5	23.16	.0001
Error	124	8998.7	72.6		

Overall, a significant difference was found ($p < .0001$); the Duncan's Multiple Range Test showed that the Surgery rotation produced significantly higher scores ($\bar{X} = 66.0$) than did Pediatrics ($\bar{X} = 51.9$) and Medicine ($\bar{X} = 48.8$) ($p < .05$). In other words, factual knowledge was not transferred from Medicine and

Pediatrics to Surgery, and Hypothesis 2 was accepted for this type of learning.

Table 10 presents the results of the analysis using interpersonal skill as the dependent variable.

TABLE 10

ANALYSIS OF COVARIANCE ON SESSION 2 PRETEST

INTERPERSONAL SKILL

Source	DF	Sum of Squares	Mean Square	F	Significance
Model	5	3333.9	666.8	9.43	.0001
Error	96	6789.3	70.7		

Overall, a significant difference was observed ($p < .0001$). However, the Duncan's Multiple Range Test revealed that there was no difference among the three disciplines, but rather a difference between session 1 and session 2 accounted for the overall significance. Based on this analysis, Hypothesis 2 was not accepted for this area of learning; that is, since there were no differences among disciplines, it cannot be said that there was no transfer of learning.

The analysis of covariance using problem solving scores as the dependent variable showed an overall significant difference (Table 11) among the three disciplines and also between the two sessions.

TABLE 11

ANALYSIS OF COVARIANCE ON SESSION 2 PRETEST

PROBLEM SOLVING

Source	DF	Sum of Squares	Mean Square	F	Significance
Model	5	29723.6	5944.7	89.1	.0001
Error	108	7207.7	66.7		

Further analysis, using the Duncan's Multiple Range Test demonstrated that the difference among disciplines was due to all

three areas being significantly different from each other, with Medicine having the highest scores ($\overline{X}$ = 59.1), followed by Surgery ($\overline{X}$ = 42.2), then Pediatrics ($\overline{X}$ = 20.5). In the area of problem solving, Hypothesis 2 was accepted; there was no transfer of learning among disciplines.

In the analysis of attitude toward health care, an overall significant difference was found ($p<.05$). These results are presented in Table 12. The Duncan's Multiple Range Test revealed that there was no significant difference among the three disciplines, but a difference did exist between session 1 and session 2.

TABLE 12

ANALYSIS OF COVARIANCE ON SESSION 2 PRETEST

ATTITUDE TOWARD HEALTH CARE

Source	DF	Sum of Squares	Mean Square	F	Significance
Model	5	420.6	84.1	3.02	.05
Error	108	30008.4	27.9		

Hypothesis 2 was not accepted in the area of attitude toward health care; there is no evidence that learning is not transferred among disciplines in this area.

The analysis of covariance using technical skill scores as the dependent variable showed no overall significant differences (see Table 13); therefore, no further analysis was done. Hypothesis 2 was not accepted for this type of learning.

TABLE 13

ANALYSIS OF COVARIANCE ON SESSION 2 PRETEST

TECHNICAL SKILL

Source	DF	Sum of Squares	Mean Square	F	Significance
Model	5	838.7	167.7	1.17	NS (.330)
Error	102	14650.0	143.6		

In summary, it was found that experience in medical and surgical disciplines enhanced factual knowledge and problem solving scores but not technical skill scores in pediatrics. Prior exposure to other disciplines did not differentially affect interpersonal skills, attitude toward health care or technical skills. Hypothesis 2 was partly confirmed.

Discussion

Previous research has tended to treat clinical training in medical education as one uniform program, ignoring the possibility of differences existing among subject areas, domains of learning, and the environment in which the instruction occurs. Clinical teaching, much more than classroom teaching, is affected by the environment in which it occurs; students are learning through active participation in a natural setting. It was decided, therefore, to first examine the environment in which clinical instruction was taking place, to generate hypotheses based on the observations which were made, and finally to test these hypotheses in a quasi-experimental design.

The first phase of the study, the ethnographic analysis of the teaching and learning environment, revealed that different disciplines (Medicine, Pediatrics and Surgery) provided quite different environments for the students. Different types of learning (factual knowledge, problem solving, technical skill, interpersonal skill, and attitude toward health care) were emphasized in each discipline; students were involved in varying activities for varying amounts of time, and the roles of the individuals involved in the teaching process (residents, interns, attending staff) differed among the rotations.

Based on these observations, it was then formally hypothesized that: 1) the different disciplines in the clerkship

program would facilitate different types of learning, and 2) there would be no significant transfer of learning among the disciplines. Both hypotheses were partially confirmed. Factual knowledge and problem solving were facilitated by different disciplines. These two types of learning were also the ones which appeared to transfer from one discipline to the other, with the remaining types (interpersonal skill, technical skill, and attitude toward health care) showing no evidence of being transferred.

These results indicate that the clinical teaching and learning process is an extremely complex one. No generalizations can be made from the study of one subject area, or one specific aspect of student learning to another. If progress is to be made in the understanding and improvement of clinical programs, educators must consider the variety of tasks that students are involved in, the number of situations in which these tasks are performed, and the types of abilities which must be used in successful performance.

The major issue addressed in this study was the use of an alternative methodology to study complex behaviour rather than the traditional paradigm. The study has given support to the theory that the ability to perform tasks is dependent on the situation in which the task is performed. Furthermore, this study supports the approach that an ethnographic (observational) analysis of complex environments must be made before measuring performance or assessing learning in such environments.

BIBLIOGRAPHY

Atkinson, P. Worlds Apart. British Journal of Medical Education, 1973, 7, 218-224.

Argyris, C. Dangers in applying results from experimental social psychology. American Psychologist, 1975, 30, 469-485.

Bashook, P. G. A conceptual framework for measuring clinical problem-solving. Journal of Medical Education. 1976, 51, 109-114.

Bateson, G. Steps to an ecology of mind. New York: Random House, 1971.

Bloom, S.W. The medical school as a social system. Milbank Memorial Fund Quarterly, 1971, 49 (2), 1-27.

Bronfenbrenner, V. Toward an experimental ecology of human development. American Psychologist, 1977, 32, 513-531.

Brunswik, E. Perception and the representative design of psychological experiments. Berkeley: University of California Press, 1956.

Brunswik, E. The conceptual framework of psychology. In International Encyclopedia of Unified Science (Vol. 1). Chicago: University of Chicago Press, 1952.

Burg, F.D.; Brownlee, R.C.; Wright, F. H.; Levine, H.; Daeschner, C. W.; Vaughan, V. C. and Anderson, J. A. A method for defining competency in Pediatrics. Journal of Medical Education, 1976, 51, 824-828.

Burg, F. D. and Lloyd, J. S. Definition of competence: A conceptual framework. In the Proceedings from the Conference on the Definitions of Competence in Specialities of Medicine. American Board of Medical Specialists, 1979.

Cole, M. and Scribner, S. Culture and thought. A psychological introduction. New York: John Wiley and Sons, 1974.

Cronbach, L. J. Beyond the two disciplines of scientific psychology. American Psychologist, 1975, 30, 116-127.

Ellsworth, P. C. From abstract ideas to concrete instances: Some guidelines for choosing natural research settings. American Psychologist, 1977, 32 604-615.

Engel, G. L. Are medical schools neglecting clinical skills? Journal of American Medical Association, 1976, 236, 861-869.

Engel, G. L. and Filling, E. M. Research approaches in health professions education -- problems and prospects. Evaluation and the Health Professions, 1981, 4 (1), 13-20.

Flanagan, J. C. The critical-requirements approach to education objectives. School Sociology, 1950, 71, 321-324.

Goetz, A.; Folse, J. R. and Peters, M. J. Chart review as a measures of clinical competence. Proceedings of the Sixteenth Annual Conference on Research in Medical Education. 1976, 40-45.

Goffman, E. Frame analysis. New York: Harper and Row, 1974.

Grayson, M.; Nugent, C. and Oken, S. L. A systematic and comprehensive approach to teaching and evaluating interpersonal skills. Journal of Medical Education, 1977, 52, 906-913.

Hammond, K. R. The social implementation of cognitive theory. In L. Petrinovich and J.L. McGaugh (Eds.), Knowing, thinking and believing. New York: Plenum Press, 1976.

Hammond, K. R. and Kern F. J. R. Teaching comprehensive medical care. Harvard University Press: Cambridge, Massachusetts, 1959.

Harden, R. M. and Gleeson, F. A. Assessment of clinical competence using an objective structured clinical examination (OSCE). Medical Education, 1979, 13, 41-54.

Hendrix, L; Carter, M. and Hintze, J. A comparison of five statistical methods for analyzing pretest-posttest designs. Journal of Experimental Education, Winter, 1978-1979, 96-102.

Hubbard, J. P.; Levit, E.J.; Schumacher, C. F. and Schnabel, T. G. An objective evaluation of clinical competence. New England Journal of Medicine, 1965, 272, 1321-1328.

Jennings, E. Linear models underlying the analysis of covariance, residual gain scores and raw gain scores. The University of Texas at Austin Measurement and Evaluation Centre.

LaDuca, A. The structure of competence in health professions. Evaluation and the health professions, 1980, 3, (3), 253-288.

Levine, H. G. and Forman, P. M. The development of cognitive knowledge in clinical clerkships. Journal of Medical Education, 1975, 50, 271-276.

McGuire, G. P. and Rutler, D. R. History taking for medical students -- differences in performance. Lancet, 1976, ii, 556-558.

McGuire, C. M.; Soloman, L. M. and Bashook, P. G. Construction and use of written simulations. New York: The Psychological Corporation, 1976.

Merton, P.; Reader, G. G. and Kendall, P. L. The student physician. Cambridge Massachusetts: Harvard University Press: 1956.

Morse, P. K.; Levy, M.; Johnson, J. C. and Bollet, A. J. Effect of previous clerkship experience on readiness for basic medical clerkship. Journal of Medical Education, 1975, 50, 199-200.

Newble, D. I. The evaluation of clinical competence. Medical Journal of Australia, 1978, 2, 180-183.

Norman, G. R. Clinical experience and the structure of memory. In the Proceedings of the Eighteenth Annual Conference on Research in Medical Education. Washington, 1979, 214-218.

O'Donohue, W. J. Evaluation of medical students during a clinical clerkship in internal medicine. Journal of Medical Education, 1978, 53, 55-58.

Patel, V. Clinical Competence of Senior Medical Students. Unpublished Ph.D. dissertation, McGill University, August, 1981.

Patel, V. A multifaceted approach to student learning during the senior clerkship. In the Proceedings of the Nineteenth Annual Conference on Research in Medical Education. Association of American Medical Colleges, Washington, 1980, 388-389.

Petrinovich, L. Molar reductionism. In L. Petrinovich and J. L. McGaugh (Eds.). Knowing, thinking and believing. New York: Plenum Press, 1976.

Prochansky, H. M. Environmental psychology and the real world. American Psychologist, 1976, 31, 303-310.

Sanagaro, P. S. and Williamson, J. W. Research in medical education -- A classification of physician performance in internal medicine. Journal of Medical Education, 1968, 43, 389-397.

Schwartz, G. F.; Gonella, J. S. and Meuduke, H. The value of pretesting in the surgical curriculum. Journal of Surgical Research, 1974, 16, 131-133.

Scott, H. M. Performance of clinical clerks on factual knowledge and problem-solving in internal medicine. Faculty of Medicine, McGill University (unpublished data).

Smiley, R. K.; Ranan, S. and Levine, D. Z. Student performance in internal medicine: The effect of previous clerkship experience. Journal of Medical Education, 1979, 54, 938-943.