journal of clinical reasoning *and* procedural competency

march 2013

VOLUME 1 NUMBER 1

Table of Contents

Thirty Years Later

In the early 1980's, I first became interested in measuring clinical reasoning while a medical student at the University of Michigan Medical School and graduate student in Education under Adrian P. Van Mondfrans, PhD, at Brigham Young University. A natural alliance was formed with several of the faculty at Michigan's Department of Medical Education, including Judith G. Calhoun, PhD, Wayne K. Davis, PhD, Fredric M.Wolf, PhD, Larry D. Gruppen, PhD and a new junior faculty member, James O. Woolliscroft, MD, who is now Dean of the University of Michigan Medical School. I was fortunate to be one of the first to research the use of virtual patient computer simulations in teaching and assessing clinical reasoning. Twenty years later, during a visit to the University of Michigan in the early 2000's, I stopped in to see Dr. Woolliscroft, then Associate Dean, and noted a copy of my dissertation on his bookshelf. Quite surprised by that, I was told that it was kept there as a reminder of what a medical student could accomplish in Medical Education. I didn't know how to respond, and was a bit embarassed that the dissertation had never been formally published. I had left Ann Arbor in 1985 to begin a busy residency, then on to an academic career helping to establish a new specialty—Emergency Medicine.

Now, 30 years later, the confidence placed in me primarily by my University of Michigan colleagues has prompted the establishment of this new Journal. For the first time, major findings from that original PhD research are being published in this Inaugural Issue of the Journal. Even though much has changed in medical education simulation research—simulations are now more realistic and have increased fidelity, with each generation becoming more "virtual" like the actual physician-patient encounter—much is still the same. We have yet to optimize the teaching and evaluation of clinical reasoning skills. We are still trying to understand how best to integrate small group study, problem-based learning, and virtual patient simulations into the medical school curriculum. Many confusing and often contradictory theories have been derived to determine exactly how novice and expert decision-makers think—and how best to teach and test that process. We continue to grapple with the concept of individual differences in learning, and whether matching learning styles and preferences to mode of instruction really matters in highly motivated medical students. Perhaps it is not too late for a three-decade old dissertation to impact Medical Education in a needed direction.

Measuring Clinical Reasoning Competency Using a Virtual Patient Model

DANE M. CHAPMAN, MD, PhD[1,2,3]
JUDITH G. CALHOUN, PhD, MBA[1]
ADRIAN P. VAN MONDFRANS, PhD[2]
WAYNE K. DAVIS, PhD[1]

Institutional Affiliations:

1. *University of Michigan Medical School, Department of Medical Education, Ann Arbor, MI*
2. *Brigham Young University, McKay School of Education, Department of Instructional Psychology and Technology, Provo, UT*
3. *University of Missouri-Columbia, School of Medicine, Center for Clinical Reasoning and Procedural Competency, Department of Emergency Medicine, Columbia, MO*

Abstract

Background:
Physicians must be thorough yet efficient in data gathering and must use decision-making strategies that limit diagnostic studies and costs, but still promote maximal diagnostic proficiency. These clinical reasoning skills are neither adequately taught nor measured in medical schools and residencies.

Objective:
To define clinical reasoning constructs *a priori* and develop clinical reasoning indices to be used with a virtual patient simulation model for teaching and assessing clinical reasoning competency.

Methods:
We used an experimental, pretest-posttest design to assess expected gains in clinical reasoning competency after three hours of virtual patient simulation practice. Computer transcripts (N=486) were generated by 81 medical students with complete data who solved one pretest, three practice, and two posttest simulations.

Results:
Four clinical reasoning constructs were identified *a priori*: proficiency, efficiency, thoroughness, and strategy, and nineteen clinical reasoning indices were defined. Multivariate ANOVA and correlational analyses revealed significant pretest-posttest differences for posttest 1 (13/19 indices) and posttest 2 (14/19 indices), supporting the instructional effectiveness of virtual patient simulation practice and the construct validity of four clinical reasoning constructs and their corresponding nineteen clinical reasoning performance indicies. Reliability (stability) and concurrent validity of indices varied with case content.

Conclusions:
Instructional effectiveness, validity and stability of four constructs and nineteen corresponding clinical reasoning indices were established for a computer-based, free-inquiry virtual patient simulation model.

Keywords:
virtual patient, simulation, clinical reasoning, clinical decision-making, competency, assessment

Introduction

High-fidelity, virtual-reality training simulations are increasingly being used for procedural training until proficiency is reached, and before allowing trainees to perform certain high-risk procedures on patients.[1-4] The Federal Drug Administration (FDA) endorsement of procedure-based simulation training is expected to cause a ripple effect throughout all of medicine.[1] Traditional methods of procedural training, including practicing upon patients, will no longer be acceptable as currently performed. While it is doubtful that the use of patients for training will ever be completely substituted with simulations, physicians will be held to higher standards of training and remediation to reduce medical errors, just as pilots have been mandated with flight simulators.[2-6]

Despite the popularity and rapid advance of procedure-based simulations in medicine, the application of cognitive-based virtual patient simulations has been noted by some experts to be stuck in time.[3-5] The "marvelous medical education machine," a complete simulator for medical education as described by Friedman,[3] has yet to be built. As its potential impact upon medical education and patient care quality is every bit as powerful as the impact of the flight simulator upon aviation, the marvelous computer will likely be built, though probably not all at once.[3-6] The ultimate virtual patient simulator will be high-fidelity—meaning it will faithfully simulate the actual physician-patient encounter. It will also be free-inquiry—meaning users can access data freely without menus or other branching limitations and without cues. Rather than text or verbal descriptions of physical exam and diagnostic test findings, actual visual and auditory responses will be provided, such as visual cues for skin rashes, cardiac and respiratory sounds, and digital images for electrocardiographs (EKGs) and radiographs. While the USMLE® step 3® computer-based case simulation exam has made notable strides in this regard, it is not the ulitmate virtual patient simulator and it still has branching and cueing limitations.[7-10]

Our aims in this study were to implement a high-fidelity, free-inquiry virtual patient simulation (VPS) model into the medical school curriculum to teach clinical reasoning (CR) skills, and then develop a scoring rubric using the VPS model as an assessment tool for measuring data-gathering and decision-making CR competencies. Specifically, we hypothesized that: (1) three hours of VPS practice with feedback would significantly impact CR competency as measured by VPS assessments, (2) CR learning constructs could be identified, and a corresponding scoring rubric of CR indicies developed to detect expected gains in CR competency, (3) Certain CR construct(s) would be case content dependent and represent "medical knowledge" and other CR constructs would be independent of any VPS case content effect, representing underlying CR "process skills", (4) stability of CR constructs (and their corresponding CR indicies) across VPS cases of varying content could be taken as a measure of reliability, (5) construct validity of CR indices would be supported if indices detected expected pretest-posttest gains (e.g., construct validity here refers to whether an index correlates with the theorized learning construct, such as "clinical reasoning proficiency," that it purports to measure), and (6) concurrent validity of CR indices would be supported if indices from the same CR construct correlated more highly than indices from different CR constructs, and the two measures were taken at the same time.[11]

Methods

Study Design:

We used an experimental pretest, posttest control group design[12] to assess expected gains in CR competency after three hours of VPS practice. To address the effects of medical information (content) upon clinical reasoning (process), pretest-posttest and practice-posttest cases of similar and dissimilar content domains were utilized as controls.

Study Setting and Population:

The study qualified for institutional review board (IRB) exemption as a curriculum innovation project and was conducted at the Taubman Health Sciences Library Learning Resource Center of the University of Michigan Medical School. Ninety-seven of 191 post-second-year medical students volunteered without compensation to participate in a computer simulation (CS) elective during a required, four-week problem-based learning curriculum (PBLC). The PBLC occurred between the preclinical and clerkship years with 23-25 CS participants being randomly assigned to each PBLC week from May 7 to June 1 after their second year.

Computer Simulation Elective:

The 6.5 hour CS elective included two sessions (3.0 and 3.5 hours) on Monday-Wednesday, Tuesday-Thursday or Wednesday-Friday mornings during which students worked through six VPSs: one 60-minute pretest (cardiology), three 60-minute practice simulations with corrective feedback (pediatric endocrinology, infectious disease and pulmonary), and two 45-minute posttests (pulmonary and cardiology). No corrective feedback was provided for pretest or posttest assessment simulations. Students were randomly assigned to work in groups of three or individually during practice simulations only. All students completed their pretest and posttest simulations as individuals.

Virtual Patient Simulations:

The multi-problem, network-based VPSs used in the study simulated the actual physician-patient encounter with high fidelity and free inquiry and included 21 patient problems among the six cases. Following an "opening scene," users assumed the role of physicians and moved to and from history, physical examination, diagnostic study, diagnosis and treatment sections without menu-driven cueing or branching limitations.[13] The VPSs were not the ultimate virtual patient, however, as artificial intelligent responses to all history, physical exam and diagnostic test inquiries were provided as text, and not virtual touch, sound or images.

Assessments and Procedure:

Computer transcripts (N=486) were generated by 81 medical students with complete data, and documented student-computer interactions for 243 hours of medical student practice and 202 hours of assessment. Outcome performance scores along nineteen predetermined CR indices (dependent variables) were derived from 243 hard-copy computer assessment transcripts (one pretest and two posttests). To standardize transcript scoring, coding regulations were developed using sample transcripts. Case-specific VPS scoring protocols provided a summary of those expert-recommended critical inquiries that had been made. Diagnosis sections were independently scored by two individuals using case-specific coding regulations that identified acceptable synonyms for diagnoses. Transcripts were scored by at least one rater who was blinded to pretest-posttest classification, and inter-rater agreement was consistently high (r>.90). While therapeutic and management plans were also computer-scored, these were ignored for the purposes of this study.

Development of Scoring Rubric:

Nineteen CR competency indices were defined *a priori* based upon a review of the medical problem-solving literature and were classified into one of four clinical reasoning constructs: proficiency, efficiency, thoroughness, and strategy (See Table 1).

Clinical reasoning proficiency referred to how effectively critical data were gathered and correct diagnoses made. The CR proficiency indices were: percent of critical data-gathering inquiries obtained for history (history proficiency), physical examination (physical examination proficiency), and diagnostic tests (diagnostic test proficiency); percent of correct diagnoses made (diagnosis proficiency); Problem Solving Index (PSI)–an

TABLE 1: Mathematical Descriptions of Nineteen Clinical Reasoning Performance Indices Derived for Use in Multi-Problem Virtual Patient Simulations

Index	Abbreviation	Description[a]
Proficiency		
History Taking	HTP	(Obtained CHT/Total CHT) X 100
Physical Examination	PEP	(Obtained CPE/Total CPE) X 100
Diagnostic Tests	DTP	(Obtained CDT/Total CDT) X 100
Correct Diagnoses	DP	(Obtained CD/Total CD) X 100
Program Solving Index	PSI	(HTP + PEP + DTP + DP) / 4
Proficiency Index	PI	(Obtained CHT + CPE + CDT) X 100 / (Total CHT + CPE + CDT)
Efficiency		
History Taking	HTE	(CHT Obtained/HTT) X 100
Physical Examination	PEE	(CPE Obtained/PET) X 100
Diagnostic Tests	DTE	(CDT Obtained/DTT) X 100
Thoroughness		
History Taking	HTT	Total HT
Physical Examination	PET	Total PE
Diagnostic Tests	DTT	Total DT
Total Data-Gathering	TDG	(HTT + PET + DTT)
Diagnosis	DT	Total D
Strategy		
History Taking	HTS	[HTT/(HTT+PET + DTT)] X 100
Physical Examination	PES	[PET/(HTT+PET + DTT)] X 100
Diagnostic Tests	DTS	[DTT/(HTT+PET + DTT)] X 100
Focused Strategy Index	FSI	(HH + PP+ DD +1) / (HP + HD + PH + PD + DH + DP + 1)
Invasiveness/Cost Index	ICI	[DTT/(HTT + PET)] X 100

[a]*Symbol Key: HT*= history taking inquiries, *PE*= physical examination inquiries, *DT*= diagnostic test inquiries, *D*= diagnoses indicated, *C*= critical inquiry or diagnosis (e.g. CHT=critical history taking inquiries), *HH*= history to history transition, *PP*= physical exam to physical exam transition, *DD*= diagnostic test to diagnostic test, *HP*= history to physical exam, HD= history to diagnostic test, *PH*= physical exam to history, *PD*= physical exam to diagnostic test, *DH*= diagnostic test to history, and *DP*= diagnostic test to physical exam transition.

average of data-gathering and decision-making proficiencies; and Proficiency Index (PI)–the percent of data-gathering critical information obtained.

Clinical reasoning efficiency was defined as the percentage of data-gathering inquiries that were critical in making the diagnosis of a patient's problem(s). Higher scores represented greater efficiency in making medical inquiries. Clinical reasoning efficiency indices included history, physical examination and diagnostic test efficiencies.

Clinical reasoning thoroughness reflected the frequency of data-gathering inquiries made or diagnoses indicated. Clinical reasoning thoroughness indices included: total number of history inquiries (history thoroughness), physical examination inquiries (physical examination thoroughness), and diagnostic test inquiries (diagnostic test thoroughness); total number of history, physical examination and diagnostic test inquiries combined (total data-gathering thoroughness); and total number of diagnoses hypothesized at the completion of each simulated case (diagnosis thoroughness).

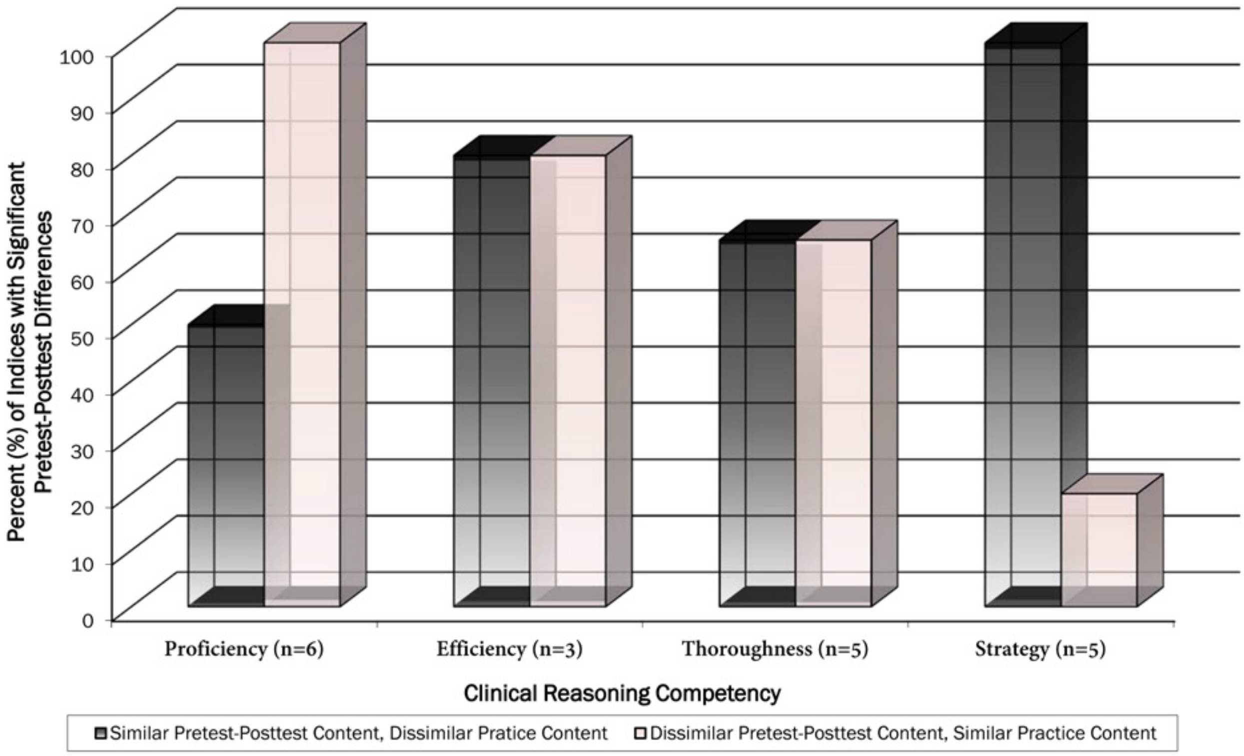
Clinical reasoning strategy referred to the cognitive strategies used to arrive at correct diagnoses. It reflected individual preference for certain data-gathering techniques (e.g. to use either a focused inquiry approach or a "shot gun" or haphazard approach). CR strategy indices included: percent of total data-gathering inquiries that relate to history taking (history strategy), physical examination (physical examination strategy), or diagnostic test (diagnostic test strategy); Focused Strategy Index—the ratio of data-gathering inquiry transitions of similar type (e.g. history to history) to all other combinations of possible inquiry transitions from one type of inquiry to another (e.g. history to physical examination, diagnostic test to history, etc), where high scores reflect a more focused and systematic data-gathering approach; and Invasiveness/Cost Index—the ratio of diagnostic test inquiries (relatively invasive and costly) to the sum of history-taking and physical examination inquiries (relatively non-invasive and less costly), where higher scores

TABLE 2: Pretest-Posttest Means (SD) for Nineteen Clinical Reasoning Indices Across Virtual Patient Simulations of Similar and Dissimilar Case Content (N=81)[a]

Index	Case 1 Pretest (Cardiology)	Case 5 Posttest (Pulmonary)	Case 6 Posttest (Cardiology)
Proficiency			
History Taking	45.1 (24.8)	56.2 (19.1)[c]	49.0 (15.0)
Physical Examination	50.9 (22.9)	67.1 (17.9)[c]	62.2 (19.5)[c]
Diagnostic Tests	58.6 (26.0)	51.7 (12.2)[b]	52.5 (16.7)[b]
Correct Diagnoses	37.0 (26.4)	46.9 (21.4)[c]	30.2 (17.5)[b]
Program Solving Index	47.9 (14.9)	55.5 (10.4)[c]	48.5 (9.7)
Proficiency Index	51.6 (16.1)	56.2 (10.8)[b]	53.1 (10.1)
Efficiency			
History Taking	8.2 (4.5)	19.3 (8.5)[c]	25.1 (10.9)[c]
Physical Examination	17.7 (8.5)	17.0 (10.2)	21.9 (10.9)[c]
Diagnostic Tests	35.4 (22.7)	43.0 (16.1)[c]	36.1 (13.6)
Thoroughness			
History Taking	23.7 (9.7)	26.1 (10.9)[b]	22.4 (10.1)
Physical Examination	12.2 (5.9)	15.0 (7.5)[c]	15.9 (6.2)[c]
Diagnostic Tests	8.4 (4.6)	9.4 (3.7)[b]	14.5 (6.4)[c]
Total Data-Gathering	44.4 (12.8)	50.5 (16.2)[c]	52.8 (14.2)[c]
Diagnosis	3.0 (1.2)	3.1 (1.1)	4.9 (1.8)[c]
Strategy			
History Taking	52.7 (12.4)	51.0 (9.9)	41.3 (13.5)[c]
Physical Examination	27.5 (8.8)	28.8 (8.6)	30.2 (9.6)[b]
Diagnostic Tests	19.8 (11.0)	20.2 (8.6)	28.4 (11.7)[c]
Focused Strategy Index	4.8 (3.3)	8.9 (5.7)[c]	9.0 (6.0)[c]
Invasiveness/Cost Index	27.6 (21.1)	26.9 (14.8)	44.3 (29.4)[c]

[a]Repeated-Measures ANOVA for Pretest - Posttest Comparisons for 81 medical students with complete data; [b]p = .050; [c]p = .010

FIGURE 1: Percent of Clinical Reasoning Indicies (N=19) with Significant Pretest-Posttest Differences Following Three Hours of Multi-Problem Virtual Patient Simulation Practice in Post-Second Year Medical Students (N=81)(ANOVA, F Test)



reflect a more invasive and costly data-gathering approach.

Data Analysis:

BMDP multivariate, factorial, repeated measures ANOVA statistics were used to determine any overall effect of three hours of VPS practice upon the four clinical reasoning constructs (proficiency, efficiency, thoroughness, and strategy), while controlling for CS weeks, group or individual practice, and a justifying strategy. CR constructs with significant multivariate effects were further defined using univariate ANOVA or ANCOVA of the individual indices of the constructs. Expected pretest-posttest gains along the nineteen CR indices were also taken as a measure of construct validity. Correlation analyses were used to determine the reliability (stability) of nineteen CR indices across cases of similar and dissimilar content and the concurrent validity of these indices in measuring one of the four CR constructs. (See also http://www.statistical-solutions-software.com/bmdp-statistical-software/bmdp/).

Results

The study sample (N=97) appeared to be representative of the entire medical school class (N=191) as VPS students did not differ significantly from other class members in ethnicity, sex, prior clinical experience on the hospital wards, or independent PBLC CR assessments (P>.050, ANCOVA). Approximately two-thirds of the VPS enrollees had never previously participated in computer-based instruction and almost one-fourth had never interacted with a computer in any capacity at the time of the original study,[13] making a selection bias, which favored students who were more comfortable with using computers for learning, unlikely.

Effect of VPS Practice:

Repeated measures factorial ANOVA analyses revealed significant pretest-posttest differences between the pretest and first posttest (13/19 indices) and between the pretest and second posttest (14/19 indices), supporting the instructional effectiveness of only three hours of VPS practice (See Table 2). It is unlikely that VPS pretest-posttest differences were a result of the PBL curriculum alone as there was no difference in PBLC CR assessments between VPS enrollees and the remainder of the medical school class during each week of the PBLC (P>.050, ANCOVA).

Effect of VPS Case Content:

Proficiency indices demonstrated pretest-posttest gains that were most notable when practice and posttest content were similar (pulmonary-pulmonary). Efficiency and thoroughness indices demonstrated significant pretest-posttest differences

TABLE 3: Correlations[a] of Nineteen Clinical Reasoning Performance Indices Across Computer Simulations of Similar and Dissimilar Case Content (N=81)[b]

Index	C1 Cardiology C5 Pulmonary	C1 Cardiology C6 Cardiology	C5 Pulmonary C6 Cardiology
Proficiency			
History Taking	.11	.23[c]	.21
Physical Examination	-.01	.23[c]	-.06
Diagnostic Tests	.10	.28[d]	-.03
Correct Diagnoses	-.03	.23[c]	.10
Program Solving Index	.03	.58[d]	.02
Proficiency Index	.11	.52[d]	.08
Efficiency			
History Taking	.06	.09	.35[d]
Physical Examination	.06	.29[d]	.19
Diagnostic Tests	.11	-.16	.24[c]
Thoroughness			
History Taking	.52[d]	.50[d]	.70[d]
Physical Examination	.44[d]	.43[d]	.63[d]
Diagnostic Tests	.42[d]	.43[d]	.63[d]
Total Data-Gathering	.53[d]	.50[d]	.70[d]
Diagnosis	.11	.31[d]	.13
Strategy			
History Taking	.35[d]	.32[d]	.54[d]
Physical Examination	.24[c]	.18	.53[d]
Diagnostic Tests	.50[d]	.40[d]	.67[d]
Focused Strategy Index	.41[d]	.35[d]	.50[d]
Invasiveness Index	.50[d]	.40[d]	.65[d]

[a]Pearson Product-Moment Correlations; [b]C1=Case 1 (Card. Pretest), C5=Case 5 (Pulm. Posttest), C6=Case 6 (Card. Posttest); [c]$p = .050$; [d]$p = .010$

regardless of case content, suggesting their stability across cases and their relation to underlying CR process skills. Strategy indices demonstrated the greatest pretest-posttest differences when posttest content was different from practice content (pulmonary-cardiology) (See Figure 1). Students became more focused in their problem-solving approach from pretest to posttest simulations as evidenced by significant improvements on the Focused Strategy Index (p=.010) regardless of case content. However, when content was unfamiliar—had not been taught during a virtual patient practice session—students used a significantly more invasive and costly problem-solving approach and relied less upon history taking and physical examination as evidenced by the Invasiveness/Cost Index (p=.010) (See Table 2).

Construct Validity, Concurrent Validity, and Strategy (Reliability):

Construct validity of the four CR constructs (proficiency, efficiency, thoroughness, and strategy) was supported by expected pretest-posttest gains after three hours of VPS practice. Concurrent validity[11] of the four CR constructs was suggested, as indices from each contruct tended to behave similarly with regard to case content and pretest-posttest effect. Higher correlations were noted among proficiency indices as expected when case content was similar (Case 1: Cardiology and Case 6: Cardiology) and were greater than pretest-postetest correlations (Case 1 and Case 5; Case 1 and Case 6; see Table 3). Concurrent validity of efficiency, thoroughness, and strategy indices was supported by generally higher correlations between the two posttests than between either posttest and the pretest (See Table 3). Concurrent validity is demonstrated when a test correlates well with a measure that has been (previously or simultaneously) validated for the same construct, or for different, but presumably related, constructs, and the two measures are taken at the same time. This is in contrast to predictive validity, where one measure occurs earlier and is meant to predict some later measure.[12] Between case correlations remained moderate to high, regardless of case content, for thoroughness and strategy indices, suggesting higher reliability (stability) of these indices across cases. Reliability of efficiency indices was less well-supported as correlations were inconsistent across case content.

Discussion

The free-inquiry VPS model, including its validated CR constructs (proficiency, efficiency, thoroughness and strategy) and nineteen CR indicies, has proven useful as both a teaching and assessment tool. We found the teaching utility of the model so profound that even as little as three hours of VPS practice resulted in significant pretest-posttest differences for many of the CR indices. This is not to say that our novice pre-clerkship students had achieved CR competency. Their mean scores remained far below expected competenecy even if defined at a 70-percent cutoff for CR proficiency indices. These results help to elucidate those aspects of CR that can be taught as process skills independent of knowledge content, and may help to resolve some of the CR teaching and assessment chaos described by Norman[14] and Elstein.[15]

Considerable CR occurs in the earliest stages of the patient presentation. Generating correct diagnostic hypotheses (i.e. hypothesis generation) has been shown to be significantly related to the patient's chief complaint and history, while physical examination and diagnostic studies contributed less to generating correct hypotheses than to eliminating alternatives (i.e. hypothesis confirmation/exclusion).[16] Moreover, students who failed to list the correct diagnosis in the differential diagnosis after obtaining the history were significantly less likely to reach the correct diagnosis at the end of the case, suggesting the critical importance of the history in medical problem solving.[17] The fact that our novice, pre-clerkship medical students had relatively low diagnosis proficiency scores compared to their data-gathering proficiencies is consistent with this finding. With their heads full of isolated, unassimilated medical facts, not organized around clinical scenarios or schemata, students did not have the key concepts or clinical features of disease patterns assimilated sufficiently to prompt their history inquiries. Still, the VPS model and CR indices were sensitive enough to detect pretest-posttest gains in both history-taking proficiency and diagnosis proficiency when content was familiar to students. These results are consistent with previous research demonstrating that medical decision-making expertise is related to one's ability to recognize content-specific disease patterns ("illness scripts") and to perceptual and cognitive skills, and that expertise is more dependent upon hypothesis generation through history taking than upon hypothesis confirmation through physical examination and diagnostic testing.[14-16]

Our results confirm that some CR skills can be enhanced or learned independent of case content, namely CR efficiency, thoroughness, and strategy. However, it is less clear which efficiency, thoroughness, or strategy adjustments would be most rewarding in terms of improved diagnostic decision making. Wolf et.al.[18] found that learning to use a competing hypothesis strategy enhanced medical problem-solving performance independent of case content.

In training clinical decision makers, medical schools and residency training programs typically emphasize thoroughness. However, the more thorough physician is not always the most expert (i.e. accurate or proficient) at clinical decision-making.[19] Increasingly, thoroughness has been taken to mean "ordering more diagnostic tests" rather than being thorough in history taking or in conducting a thorough physical examination. David Sklar,[20] in his editorial "Beginning the Journey" as the new editor-in-chief of Academic Medicine, has noted that "CT scans and ultrasounds have virtually replaced the traditional physical examination, and computers have invaded the consultation room, interposing themselves between the clinician and the patient, diverting the clinician's attention from conversations with the patient to the documentation requirements demanded by payers and employers." This is a worrisome trend that threatens our professional identity as health care providers. The relationship "between the healer and the sick, the most sacred, core responsibility and privilege in medicine" is being threatened.[20]

In our attempt to teach and assess core competencies through VPSs, we must be on guard not to lose the sacred trust of our patients. It seems contradictory to teach physician-patient interactions using computer-based technologies that may be the very cause of our eroding physician-patient relationships. However, if properly designed, VPS could be useful in teaching and assessing professionalism and the other core competencies identified by the Accreditation Council on Graduate Medical Education (ACGME).[21] The VPS model and CR indices could also be implemented to study intervention effects upon CR competency. It has been suggested that decision making could be enhanced and its teaching facilitated if disease-specific, data-gathering elements were identified and characterized as most consistent and predictive of each competing diagnostic hypothesis. Understanding the optimal disease-differentiating pivotal elements, key concepts, features[14-16] and knowledge structures[14] would seem to significantly augment acquisition of clinical reasoning skills—especially when programmed into virtual patient simulators.[3-5, 22-23] Developers of newer generation, virtual patient simulators would also do well to incorporate the free-inquiry approach, without cueing or branching limitations. Such cognitive-based simulators would also be most useful if they incorporated an artificial intelligence function that responded to user treatments in disease-predictable ways, such that users are able to perform "what if" inquiries as they learn.[3-5, 22-24]

This study has limitations. It was conducted nearly three decades ago as part of a PhD dissertation,[13] and was never formally published. With recent developments in the ACGME core competencies,[21] new accreditation system (NAS) and milestones,[25] the *a priori* development and validation of CR constructs with a scoring rubric using free-inquiry VPSs has greater relevance now than thirty years ago. VPSs have changed in some ways that might impact study results. However, one could argue that the free-inquiry capability of the VPS model in this study is the gold standard which has yet to be achieved by the USMLE® step 3® computer-based or OSCE-based exam.[5-10] Further study is needed to apply generalizability analysis of the scoring rubric to better understand inter-case variability. Generalizability refers to external validity and is limited when the cause or independent variable (e.g., three hours of VPS practice) is influenced by other factors—all threats to external validity or generalizability interact with the independent variable.[26] Although this study was conducted at a single institution at a single point in time some years ago, more than half of a large medical school class participated, and the results of this study would be expected to generalize to other post-second year medical students with similar aptitudes and experiences. It is less clear whether results would generalize to medical students in their clinical years or to residents and physicians.

In summary, four clinical reasoning constructs of proficiency, efficiency, thoroughness and strategy were defined *a priori* and validated using a high-fidelity, free-inqiry, computer-based virtual patient simulation model. With ever-changing protocols and increasing medical knowledge, VPS may be helpful in positioning medical students and trainees for life-long learning as part of their daily clinical practice.[21, 24-25] If the ultimate goal for incorporating VPSs into all levels of medical education is to promote improved quality of care for patients,[1- 2] while regaining a new sense of commitment to the clinician-patient relationship,[20] then we will ultimately succeed in building the marvelous medical education machine. After thirty years of processing and assimilation, the VPS machine may be capable of both teaching CR skills and producing a scoring rubric that can detect subtle differences in clinical data-gathering and decision-making core competencies.

Received for publication: August 20, 2012. Revisions received: February 20, 2013. Accepted for publication: February 20, 2013. Available online: February 21, 2013.

Presented in part at the Society for Academic Emergency Medicine (SAEM) Annual Meeting, San Antonio, Texas, May 21-24, 1995; A National Symposium on Microcomputers in Health Care Education, Omaha, Nebraska, April, 1985.

Contributions of Authors: Conception (DMC, JGC, APV, WKD), design (DMC, JGC), analysis and interpretation (DMC, JGC), drafting (DMC, JGC, WKD), and revising (DMC) the article. DMC takes responsibility for the paper as a whole.

Acknowledgements: The authors are indebted to Robert H. Bartlett, MD and Richard G. Judge, MD, course directors of the Problem-Based Learning Curriculum, for their enthusiastic endorsement of the medical problem solving computer simulation elective; the ninety-seven University of Michigan medical students who enrolled in the elective; the Taubman Medical Library Learning Resource Center staff for their essential role in making the computer simulation elective an overwhelming success, namely: Cindy A. McBride, Patricia Tomlin, Ellen S. Hoffman, Achla B. Karnani, Chris M. Chapman, Denise M. Axelrod, Patricia W. Martin, Dennis M. Poupart, Jackie E. McKenzie, Lois L. Katon, Mark D. Heidt, William A. Vicini, and Sharon J. Love; and, Fredric M. Wolf, PhD, Elaine M. Hockman, PhD, Nancy P. Allen, PhD, C. Victor Bunderson, PhD, Russell T. Osguthorpe, PhD and Sally H. Cavanaugh, PhD for their insightful review and recommendations to improve this work; and, to Becky Bluett and Christine Downs for their assistance in preparing the manuscript.

Funding: In part by a University of Michigan curriculum innovation grant.

Conflicts of Interest: No personal, commercial, political, academic or financial conflict of interest was reported by any of the authors. DMC is editor-in-chief of JCRPC.

Ethical Approval: Authors report the study qualified for IRB exemption.

Correspondence: Dane M. Chapman, MD, PhD, Director, Center for Clinical Reasoning and Procedural Competency, Department of Emergency Medicine, University of Missouri-Columbia School of Medicine, 1 Hospital Drive, Columbia, MO, 65201, e-mail: <chapmandan@health.missouri.edu>

References

1. Gallagher AG, Cates CU. Approval of virtual reality training for carotid stenting: what this means for procedure-based medicine. *JAMA* 2004 Dec 22;292(24):3024-26.
2. AAMC Report of the Ad Hoc Committee of Deans. Educating doctors to provide high quality medical care. 2004: 1-12. Available from: https://members.aamc.org/eweb/upload/Educating%20Doctors%20to%20Provide%20July%202004.pdf.
3. Friedman CP. The marvelous medical education machine or how medical education can be unstuck in time. *Acad Med.* 2000;75(10):S137-142.
4. Cook DA, Triola MM. Virtual patients: A critical literature review and proposed next steps. *Med Educ*. 2009;43(4):303-11.
5. Courteille O, Bergin R, Stockeld D, Ponzer S, Fors U. The use of a virtual patient case in an OSCE-based exam—a pilot study. *Med Teach.* 2008;30(3):e66-76.
6. Noble C. The relationship between fidelity and learning in aviation training and assessment. *J of Air Transport.* 2002 7(3):33-54.
7. Feinberg RA, Swygert KA, Haist SA, Dillon GF, Murray CT. The impact of postgraduate training on USMLE® step 3® and its computer-based case simulation component. *J Gen Intern Med.* 2012 Jan;27(1):65-70.
8. Harik P, Cuddy MM, et al. Assessing potentially dangerous medical actions with the computer-based case simulation portion of the USMLE step 3 examination. *Acad Med.* 2009 Oct;84(10 Suppl):S79-82.
9. Andriole DA, Jeffe DB, Hageman HL, Whelan AJ. What predicts USMLE Step 3 performance? *Acad Med.* 2005 Oct;80(10 Suppl):S21-4.
10. Margolis MJ, Clauser BE, Harik P. Scoring the computer-based case simulation component of USMLE Step 3: a comparison of preoperational and operational data. *Acad Med.* 2004 Oct;79(10 Suppl):S62-4.
11. McIntire SA and Miller LA, *Foundations of Psychological Testing*, 2nd Edition, Thousand Oaks, CA:Sage Publishing Co., 2005.
12. Campbell DT, Stanley JC 1963. Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally College Publishing Company.
13. Chapman DM. Teaching and evaluating clinical reasoning through computer-based patient management simulations. *Dissertation Abstracts International.* 1985;46:784-B.
14. Norman GR. The epistemology of clinical reasoning. *Acad Med.* 2002;75-S127-133.

15. Elstein AS. Clinical problem solving and decision psychology: Comment on "The epistemology of clinical reasoning." *Acad Med*. 2002;75(10):S134-36.

16. Gruppen LD, Woolliscroft JO, Wolf FM. The contribution of different components of the clinical encounter in generating and eliminating diagnostic hypotheses. *Proc Annu Conf Res Med Educ.* 1988;27:242-7.

17. Gruppen LD, Palchik NS, Wolf FM, et al. Medical student use of history and physical information in diagnostic reasoning. *Arthritis Care Res*. 1993;6(2):64-70.

18. Wolf FM, Gruppen LD, Billi JE. Use of the competing-hypotheses heuristic to reduce "pseudodiagnosticity'. *J Med Educ.* 1988 Jul;63(7):548-54.

19. Voytovich AE, Rippey RM, Copertino L. Scorable problem lists as measures of clinical judgment. *Eval Health Prof.* 1980;3:159-171.

20. Sklar D. Beginning the journey. *Acad Med.* 2013;88(1):1-2.

21. Chapman DM, Hayden S, Sanders AB et al. Integrating the Accreditation Council for Graduate Medical Education core competencies into the Model of the Clinical Practice of Emergency Medicine. *Ann Emerg Med.* 2004;33(6):756-769.

22. Stevens SM, Goldsmith TE, Summers KL et. al. Virtual reality training improves students' knowledge structures of medical concepts. *Medicine Meets Virtual Reality* 13, James D. Westwood et al. (Eds.), Amsterdam:IOS Press, 2005, pp. 519-525.

23. Deterding R, Milliron C, Hubal R. The virtual pediatric standardized patient application: formative evaluation findings. *Medicine Meets Virtual Reality* 13, James D. Westwood et al. (Eds.), Amsterdam:IOS Press, 2005, pp. 105-107.

24. Jarrell BE. Simulation for teaching decision making in medicine: The next step. Presented at Medicine Meets Virtual Reality 13, Long Beach, CA, Jan 28, 2005.

25. Nasca TJ, Philibert I, Brigham T, Flynn TC. The next GME accreditation system — Rationale and benefits. NEJM Special Report. http://www.acgme-nas.org/assets/pdf/NEJMfinal.pdf

26. Shadish W, Cook T, Campbell D. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston:Houghton Mifflin, 2002.

Acquiring Clinical Reasoning Competency: Group versus Individual Practice of Virtual Patient Simulations

DANE M. CHAPMAN, MD, PhD[1,2,3]
JUDITH G. CALHOUN, PhD, MBA[1]
ADRIAN P. VAN MONDFRANS, PhD[2]
WAYNE K. DAVIS, PhD[1]

Institutional Affiliations:

1. *University of Michigan Medical School, Department of Medical Education, Ann Arbor, MI*
2. *Brigham Young University, McKay School of Education, Department of Instructional Psychology and Technology, Provo, UT*
3. *University of Missouri-Columbia, School of Medicine, Center for Clinical Reasoning and Procedural Competency, Department of Emergency Medicine, Columbia, MO*

Abstract

Background:
It is unknown whether group or individual practice using free-inquiry virtual patient simulations would most facilitate acquisition of clinical reasoning skills required of competent physicians.

Objective:
To determine the effect of virtual patient simulation group practice on clinical reasoning competency.

Methods:
We used an experimental, pretest-posttest, control group design. Ninety-seven of 191 post-second year medical students were randomly assigned to group practice or individual practice and solved six virtual patient simulations: one pretest (individual), three practice (group or individual), and two posttest (individual) simulations. Multivariate ANOVA and univariate ANCOVA statistics were used to compare groups.

Results:
Computer transcripts (N=486) were generated by 81 post-second year medical students with complete data. Compared to individual-practice students (n=41), group-practice students (n=40) performed as well as or better on 18 of 19 clinical reasoning proficiency measures, demonstrated greater overall clinical reasoning proficiency, indicated more diagnostic hypotheses and used more focused inquiries. Individual-practice students were more efficient in making critical physical examination inquiries.

Conclusions:
Instructional effectiveness was established for both individual and group virtual patient simulation practice, with a combined group practice and virtual patient simulation effect in promoting clinical reasoning competency.

Keywords:
virtual patient simulation, group practice, clinical reasoning, clinical decision-making, medical students

Introduction

In our current healthcare climate, physicians are rewarded for being sufficiently thorough, yet efficient in data gathering and for using problem-solving strategies that limit diagnostic tests and commensurate costs, but still promote maximal diagnostic proficiency. These critical clinical reasoning (CR) skills (proficiency, efficiency, thoroughness, and strategy) are neither adequately mastered nor measured in medical schools and residency programs.[1-3] CR involves both data gathering (i.e. history taking, physical examination, and the selection and interpretation of appropriate diagnostic tests) and diagnostic and therapeutic decision-making skills.[3-4] Cognitive models are needed to teach and assess data-gathering and decision-making competencies,[5-6] preferably earlier in medical school since the organization of clinical knowledge and the directionality of CR acquired during medical school carries over into subsequent resident and physician performance.[5-9] Knowledge organization and schema acquisition seem to be more important for CR expertise than the use of problem-solving methods, and their development over many years of experience requires exposure to many patient cases while emphasizing the association of disease-specific features, signs and symptoms.[7,9]

High-fidelity, free-inquiry virtual patient simulations (VPS) provide increasingly sophisticated opportunities to engage in virtual patient encounters, and have been implicated in teaching and testing cognitive CR skills.[10-13] Virtual patient simulations are considered "high-fidelity" if they closely simulate an actual physician-patient encounter, and are "free-inquiry" if they allow free questioning without menu or branching limitations. It seems likely that such simulations could provide trainees with the equivalent of "many years of patient care experience," and facilitate their knowledge organization and schema acquisition.[3,11-12]

While the cost of high-fidelity VPSs has limited their widespread use in training,[10] group study could facilitate their integration into medical school and residency curricula by requiring fewer start-up multimedia configurations. Small group study using the problem-based learning (PBL) format has been shown to improve CR competency as measured by United States Medical Licensing Examination (USMLE) and residency program director evaluations, but with greater financial and faculty resource costs.[8] Once developed, VPSs could markedly decrease dependance upon overburdened faculty and limited training resources.[13] However, it is unknown whether GP would facilitate or hinder CR acquisition using an interactive, free-inquiry VPS model. It is conceivable that VPS GP might dilute the frequency of individual student-simulation interactions resulting in an antagonistic rather than a synergistic effect. Accordingly, we conducted this study to determine whether free-inquiry VPS group (GP) versus individual (IP) practice would impact the development of CR competency. We hypothesized that: (1) VPS practice would enhance medical student CR competency, (2) CR competency would vary by VPS case content, and (3) GP students would do as well as or better than IP students on measures of CR competency.

Methods

Study Design:

Using an experimental, pretest-posttest, control group design, we assessed the effect of three hours of VPS GP versus IP practice (independent variable) upon nineteen previously established CR dependent outcome measures.[3]

Study Setting and Population:

The study qualified for internal review board (IRB) exemption as a curriculum innovation project and was conducted at the Taubman Health Sciences Library Learning Resource Center at the University of Michigan Medical School. Ninety-seven of 191 pre-clinical medical students participated in a one-week VPS elective during a required four-week PBL curriculum, with 23-25 participants being randomly assigned to each of the four weeks between the preclinical and clerkship years.

Virtual Patient Simulation Elective:

The 6.5 hour VPS elective included two morning sessions (3.0 and 3.5 hours) during which students worked through six free-inquiry virtual patient simulations: one 60-minute pretest as individuals (Case 1: Cardiology), three 60-minute GP or IP simulations with corrective feedback (Case 2: Pediatric Endocrinology, Case 3: Infectious Disease, and Case 4: Pulmonary), and two 45-minute posttests as individuals (Case 5: Pulmonary and Case 6: Cardiology).

Group Versus Individual Practice:

GP students worked in groups of three and were assigned to one of three roles which changed until each group member had experienced each role: (1) typist: typed group inquiries at the keyboard; (2) recorder: recorded times, type of inquiries, diagnostic hypotheses and likelihood rankings on a VPS log; and (3) chairperson: insured that all group members participated in making decisions, and cast the deciding vote if group members were indecisive. IP students worked through the three practice VPS alone and were responsible for typing inquiries at the keyboard and maintaining their own VPS log.

Virtual Patient Simulations:

Multi-problem, text-driven, network-based virtual patient simulations were selected for practice and assessment as they were the most sophisticated high-fidelity, free-inquiry simulations available at the time of the study.[3] Following an "opening scene," users assumed the role of physicians and moved to and from history, physical examination, diagnostic study, diagnosis and treatment sections without cueing or branching limitations. The VPS responded to user inquiries with questions or feedback in predictable ways using artificial intelligence.

Assessments and Procedure:

VPS transcripts (N=486) were generated by 81 medical students with complete data and documented student-simulation interactions for 243 hours of medical student practice, and 202 hours of assessment. Outcome performance scores along the nineteen predetermined CR indices were derived from 243 hard-copy VPS assessment transcripts (one pretest and two posttests) and were classified into one of four previously validated[3] clinical reasoning constructs: proficiency, efficiency, thoroughness, and strategy (See Table 1).

Clinical reasoning proficiency refers to how effectively critical data were gathered and correct diagnoses made. The dependent CR proficiency variables were: percent of critical data gathering inquiries obtained (history, exam, and diagnostic test proficiencies), percent of correct diagnoses made (diagnosis

TABLE 1: Mathematical Descriptions of Nineteen Clinical Reasoning Performance Indices Derived for Use in Multi-Problem Virtual Patient Simulations

Index	Abbreviation	Description[a]
Proficiency		
History Taking	HTP	(Obtained CHT/Total CHT) X 100
Physical Examination	PEP	(Obtained CPE/Total CPE) X 100
Diagnostic Tests	DTP	(Obtained CDT/Total CDT) X 100
Correct Diagnoses	DP	(Obtained CD/Total CD) X 100
Program Solving Index	PSI	(HTP + PEP + DTP + DP) / 4
Proficiency Index	PI	(Obtained CHT + CPE + CDT) X 100 / (Total CHT + CPE + CDT)
Efficiency		
History Taking	HTE	(CHT Obtained/HTT) X 100
Physical Examination	PEE	(CPE Obtained/PET) X 100
Diagnostic Tests	DTE	(CDT Obtained/DTT) X 100
Thoroughness		
History Taking	HTT	Total HT
Physical Examination	PET	Total PE
Diagnostic Tests	DTT	Total DT
Total Data-Gathering	TDG	(HTT + PET + DTT)
Diagnosis	DT	Total D
Strategy		
History Taking	HTS	[HTT/(HTT+PET + DTT)] X 100
Physical Examination	PES	[PET/(HTT+PET + DTT)] X 100
Diagnostic Tests	DTS	[DTT/(HTT+PET + DTT)] X 100
Focused Strategy Index	FSI	(HH + PP+ DD +1) / (HP + HD + PH + PD + DH + DP + 1)
Invasiveness/Cost Index	ICI	[DTT/(HTT + PET)] X 100

[a]*Symbol Key: HT*= history taking inquiries, *PE*= physical examination inquiries, *DT*= diagnostic test inquiries, *D*= diagnoses indicated, *C*= critical inquiry or diagnosis (e.g. CHT=critical history taking inquiries), *HH*= history to history transition, *PP*= physical exam to physical exam transition, *DD*= diagnostic test to diagnostic test, *HP*= history to physical exam, HD= history to diagnostic test, *PH*= physical exam to history, *PD*= physical exam to diagnostic test, *DH*= diagnostic test to history, and *DP*= diagnostic test to physical exam transition.

proficiency), Problem Solving Index–an average of data gathering and decision making proficiencies, and Proficiency Index–the percent of data-gathering critical information obtained.

Clinical reasoning efficiency was defined as the percentage of data gathering inquiries that were critical in making the diagnosis of a patient's problem(s). Higher scores represent greater efficiency in making medical inquiries. Clinical reasoning efficiency variables included history, physical examination and diagnostic test efficiencies.

Clinical reasoning thoroughness reflects the frequency of data gathering inquiries made or diagnoses indicated. Clinical reasoning thoroughness variables included: total number of history inquires (history thoroughness), physical examination inquiries (physical examination thoroughness), and diagnostic test inquiries (diagnostic test thoroughness); total number of history, physical examination and diagnostic test inquiries combined (total data gathering thoroughness); and total number of diagnoses hypothesized at the completion of each simulated case (diagnosis thoroughness).

Clinical reasoning strategy refers to the cognitive strategies used to arrive at correct diagnoses; and, reflects individual preference for certain data gathering techniques (e.g. to use either a focused inquiry approach or a "shot gun" or haphazard approach). CR strategy indices included: percent of total data-gathering inquiries that relate to history taking (history strategy), physical examination (physical examination strategy), or diagnostic test (diagnostic test strategy); Focused Strategy Index–the standardized proportion of data gathering inquiry transitions of similar type (e.g. history to history) to all other combinations of possible inquiry transitions from one type of inquiry to another (e.g. history to physical examination, diagnostic test to history, etc), where high scores reflect a more focused and systematic data-gathering approach; and Invasiveness/Cost Index–the standardized proportion of diagnostic test inquiries (relatively invasive and costly) to the sum of history taking and physical examination inquiries (relatively non-invasive and less costly), where higher scores reflect a more invasive and costly data-gathering approach.

Data Analysis:

BMDP[14] multivariate factorial, repeated-measures ANOVA statistics were used to determine any overall effect of GP/IP practice (independent variable) upon the four CR performance constructs. When pretest differences were found between treatment groups, then analyses of covariance (ANCOVA) were utilized to adjust for pretreatment differences with the pretest treated as the covariate. If a significant multivariate effect was observed foir a CR construct, then univariate ANOVA or ANCOVA statistics were used to test for CR index differences among treatment groups.

Results

The study sample (N=97) appeared to be representative of the entire medical school class (N=191) as VPS students did not differ significantly from other class members on ethnicity, sex, prior clinical experience on the hospital wards, or on PBL CR assessments. Complete hard-copy VPS assessment transcripts (n=243) were obtained for 81 study participants: GP(n=40) and IP(n=41).

VPS Practice. Significant pretest-posttest differences were found for both treatment groups (GP/IP) and suggest the utility of three hours of high-fidelity VPS practice in teaching selective CR skills (See Table 2).

VPS Case Content. Multivariate analyses of posttests (Case 5

TABLE 2: Group Practice (GP) and Individual Practice (IP) Pretest and Posttest Mean Scores (SD) for Nineteen Clinical Reasoning Performance Measures[a]

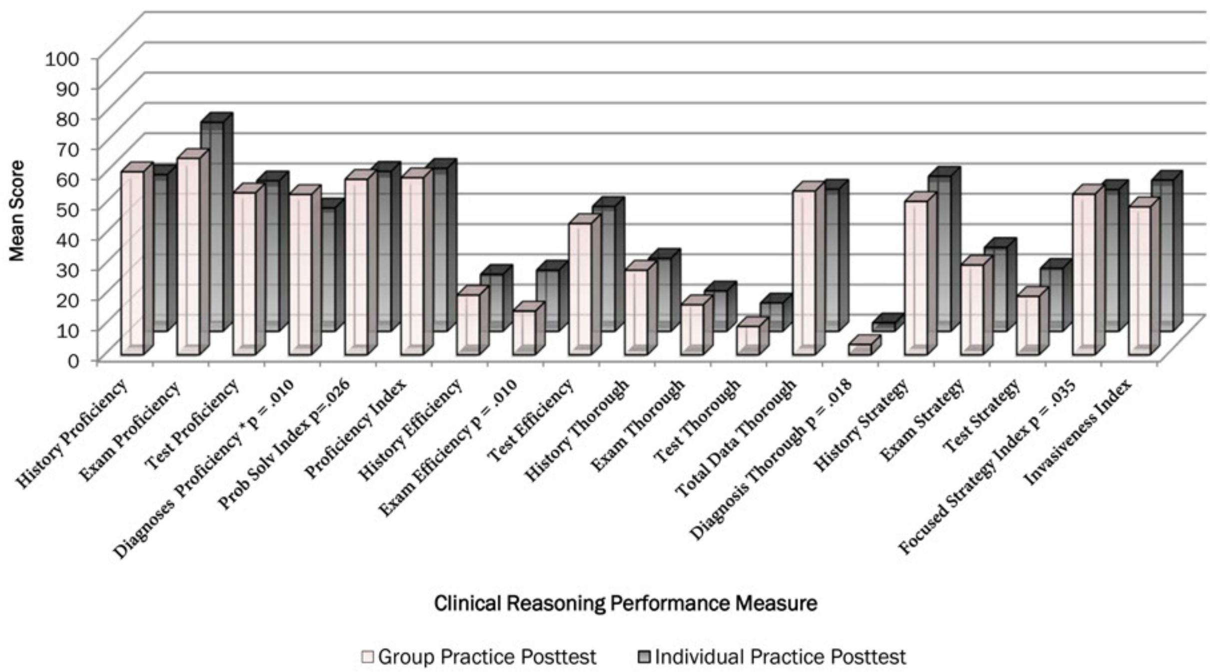
Index	Case 1 Pretest (Cardiology)		Case 5 Posttest (Pulmonary)		Case 6 Posttest (Cardiology)	
	GP	IP	GP	IP	GP	IP
Proficiency						
History Taking	50.0 (24.9)	40.2 (23.0)	60.6 (15.6)[c]	51.8 (21.2)[c]	51.0 (16.5)	47.1 (13.5)
Physical Examination	57.5 (22.1)	44.5 (22.0)	65.0 (18.4)[c]	69.1 (17.3)[c]	62.0 (19.1)[c]	62.4 (20.0)[c]
Diagnostic Tests	62.5 (24.7)	54.9 (26.9)	53.6 (13.2)[c]	49.8 (11.1)[c]	53.6 (15.5)[c]	51.5 (17.9)[c]
Correct Diagnoses	38.3 (28.8)	35.8 (24.0)	53.1 (22.8)[b,c]	40.8 (18.3)[b,c]	32.5 (18.0)[c]	28.0 (17.0)[c]
Program Solving Index	52.1 (15.0)	43.9 (13.9)	58.1 (9.3)[b,c]	52.9 (10.9)[b,c]	49.8 (9.4)	47.3 (9.9)
Proficiency Index	56.7 (15.8)	46.5 (15.0)	58.6 (9.2)[c]	53.9 (11.8)[c]	54.3 (10.1)	51.9 (10.4)
Efficiency						
History Taking	7.8 (4.0)	8.5 (5.8)	19.8 (8.1)[c]	18.9 (8.9)[c]	24.5 (11.1)[c]	25.7 (10.7)[c]
Physical Examination	17.9 (9.6)	17.5 (7.4)	14.5 (7.9)[b]	20.3 (11.4)[b]	21.0 (8.8)[c]	22.8 (10.3)[c]
Diagnostic Tests	44.8 (25.4)	26.3 (15.0)	43.5 (13.8)[c]	41.3 (18.2)[c]	36.5 (9.5)	35.6 (16.7)
Thoroughness						
History Taking	26.4 (10.2)	21.0 (8.5)	28.1 (12.4)[c]	24.2 (9.1)[c]	23.6 (11.1)	21.2 (9.0)
Physical Examination	13.6 (5.3)	10.9 (5.4)	16.6 (8.0)[c]	13.4 (6.6)[c]	16.3 (6.1)[c]	15.4 (6.4)[c]
Diagnostic Tests	7.0 (3.8)	9.8 (4.9)	9.4 (3.4)[c]	9.4 (4.0)[c]	13.8 (4.5)[c]	15.2 (7.8)[c]
Total Data-Gathering	47.1 (12.7)	41.7 (12.4)	54.2 (16.6)[c]	47.0 (15.2)[c]	53.8 (12.7)[c]	51.8 (15.6)[c]
Diagnosis	3.0 (1.2)	3.1 (1.1)	3.4 (1.2)[b]	2.8 (1.0)[b]	4.6 (1.5)[c]	5.1 (2.1)[c]
Strategy						
History Taking	55.7 (12.4)	49.7 (11.8)	50.8 (10.6)	51.3 (9.3)	42.0 (13.9)[c]	40.6 (13.3)[c]
Physical Examination	28.9 (8.2)	26.1 (9.3)	29.8 (8.3)	27.8 (8.8)	30.4 (9.4)[c]	30.1 (10.0)[c]
Diagnostic Tests	15.4 (8.9)	24.2 (11.2)	19.4 (9.5)	20.9 (7.7)	27.5 (12.1)[c]	29.3 (11.4)[c]
Focused Strategy Index	53.3 (12.2)	48.3 (7.6)	53.2 (12.3)[b]	46.9 (6.2)[b]	51.4 (9.0)	48.3 (10.7)
Invasiveness Index	46.0 (6.5)	53.1 (10.8)	49.0 (9.3)	50.0 (7.4)	49.6 (10.2)	50.7 (9.9)

[a]Repeated-Measures, Factorial ANCOVA for Pretest (Case 1) and Posttest (Case 5 and Case 6) comparisons for 81 medical students (GP, n=40; IP, n=41) with complete data.

[b]ANCOVA, F-Test, $p < .050$, for significant GP (n=40) versus IP (n=41) differences.

[c]ANCOVA, F-Test, $p < .050$, for significant pretest-posttest differences (N=81).

FIGURE 1: Mean Scores on Nineteen Case 5 Clinical Reasoning Measures after Three Hours of Group or Individual Practice Using Virtual Patient Simulations in Post-Second Year Medical Students (N=81) (*ANCOVA, F test)



and Case 6) revealed significant GP/IP differences only for the posttest with content similar to that encountered during a practice simulation (Case 5), and were related to CR thoroughness (Wilks' λ=12.1, F(4, 62)=2.89; p =.03), and strategy constructs (Wilks' λ=14.2, F(5, 61)=2.67; p = .03), although differences between treatment groups approached significance along the efficiency construct as well (Wilks' λ=8.40, F(3, 63)=2.71; p = .052) (See Table 2). There were no multivariate GP/IP differences on the posttest (Case 6) with content not previously encountered on a practice simulation.

Group Versus Individual Practice. GP students performed as well as or better than IP students on 18 of 19 CR competency measures (See Table 2; Figure 1). Despite randomization into GP/IP treatment groups, multivariate analyses revealed pretest GP/IP differences across clinical reasoning proficiency (Wilks' λ=15.0, F(6, 60)=2.31; p = .04), efficiency (Wilks' λ=14.6, F(3, 63)=4.71; p = .00), thoroughness (Wilks' λ=16.5, F(4, 62)=3.93; p = .01), and strategy measures (Wilks' λ=16.4, F(5, 61)=3.08; p = .02).

When covarying on the pretest, univariate ANCOVA of the Case 5 posttest detected significantly higher GP versus IP proficiency scores along Diagnosis Proficiency (F(1, 76)=7.06, p = .010) and the Problem-solving Index (F(1, 76)=5.17, p = .026), but not along any of the indices measuring only data gathering proficiency. Univariate ANCOVA of Case 5 efficiency scores detected a significant decrease in GP physical examination efficiency compared to IP students (F(1, 76)=7.03, p = .010). Significant differences in GP versus IP Case 5 thoroughness scores were found only for Diagnosis Thoroughness (F(1, 76)=5.85, p = .018), but not for any of the data gathering thoroughness measures. Univariate ANCOVA of Case 5 strategy measures revealed a significant tendency for GP students to use a more focused, and less haphazard data gathering strategy compared to IP students (F(1, 76)=4.59, p = .035) (See Figure 1). When covarying on the pretest, no significant GP/IP treatment effects were observed for the second posttest (Case 6) with content not previously encountered on a practice simulation (See Table 2).

Discussion

The results of this study confirm the efficacy of both group and individual practice in teaching CR skills and demonstrate that free-inquiry VPSs can be successfully implemented into the medical school curriculum. Both GP and IP students acquired selective CR skills after only three hours of free-inquiry VPS practice. That GP is at least as effective as IP in teaching all but physical exam efficiency, and superior to IP in teaching more focused data gathering and more elaborate hypothesis generation are important findings. As expected, GP/IP treatment differences were found only with case content encountered during VPS practice.

GP students acquired CR skills beyond that expected from interacting with the VPS alone. GP students were able to use their collective knowledge on a practice case of similar content, which carried over to their more focused data gathering and to greater diagnostic hypotheses generated as individuals on the assessment case of similar content (Case 5 Posttest: Pulmonary). Presumably as a result of generating more hypotheses,[15] GP students were also more likely to make correct diagnoses; and this may be the single greatest impact of VPS GP upon acquisition of CR competency.

Group use of VPSs would greatly reduce initial VPS purchasing and upkeep costs when implementing VPSs into the medical school curriculum. The challenge of doing more with less, with greater training expectations and reduced training resources, has greatly impacted health professions education.[13,16] Combining virtual patients and small group study using computer-based clinical scenarios, web-based and otherwise may help to reconcile this seeming paradox of better training with fewer resources.[13,17,18]

Since students were randomly assigned to treatment groups, it is unlikely that GP/IP pretest differences along all four CR constructs were due to chance alone. What, then could account for these differences? One tenable explanation is that a GP/IP treatment effect occurred prior to the pretest. An indirect GP treatment could have occurred as students were informed of their GP/IP treatment assignments several days before the pretest assessment. They were not, however, given the identity of the other group members. It is possible that students assigned to the GP treatment were more motivated to learn CR skills in anticipation of performing with peers.

This study has limitations. It was conducted nearly three decdes ago as part of a PhD dissertation,[19] and was never formally published. While study findings have become relevant with developments in VPS training and assessment, VPSs have changed in some ways that might impact study results. Each generation of VPSs have become more sophisticated with patient scenario video clips, actual EKGs and radiographs requiring user interpretation, and more advanced scoring, data storage, retrieval and web-based capabilities. In a study of web-based VPSs, users found demonstrations of physical exam abnormalities in heart or lung sounds, skin lesions, and neurological findings quite helpful.[20] Still, similar or more concerning limitations exist in OSCE-based and USMLE® step 3® computer-based exams and in live simulated patients.[2] The essential VPS components required for teaching and assessing clinical reasoning have remained the same: free inquiry, high fidelity, no cueing or branching limitations, and artificial intelligence interactive capabilities that require users to indicate history, exam and laboratory inquiries. Finally, the study was conducted at a single traditional medical school with post-second-year, pre-clinical medical students. It is uncertain how the results would generalize to clinical medical students, residents, or physicians in continuing medical education. In this regard, web-based VPSs have been found to have greater acceptance among pre-clinical second-year medical students compared to clinical fourth-year students.[20]

In summary, pre-clinical medical students assigned to VPS GP performed as well as or better than those assigned to VPS IP on 18 of 19 CR competency measures. This is an important finding since GP requires fewer VPS training resources, and would thereby facilitate implementing virtual patient simulations into the medical school curriculum. Moreover, these results suggest that VPS and GP had a combined or additive facilitating effect upon student acquisition of CR skills. The indirect treatment effect of prior knowledge of being assigned to work with peers is a new finding and has many implications for motivating students to acquire CR competency.

Received for publication August 13, 2012. Revisions received January 15, 2013. Accepted January 19, 2013. Available online February 1, 2013.

Presented in part as an "Innovations in Medical Education Exhibit" at the 108th annual meeting of the Association of American Medical Colleges (AAMC), Washington, DC, November 2-4, 1997; and at the 27th annual meeting of the Society for Academic Emergency Medicine (SAEM), Washington DC, May 22, 1997.

Contributions of Authors: Conception (DMC, JGC, APV, WKD), design (DMC, JGC), analysis and interpretation (DMC, JGC), drafting (DMC, JGC, WKD), and revising (DMC) the article. DMC takes responsibility for the paper as a whole.

Acknowledgements: The authors are indebted to Robert H. Bartlett, MD and Richard D. Judge, MD, course directors of the Problem-Based Learning Curriculum, for their enthusiastic endorsement of the medical problem solving computer simulation elective; the ninety-seven University of Michigan medical students who enrolled in the elective; the Taubman Medical Library Learning Resource Center staff for their essential role in making the computer simulation elective an overwhelming success; Fredric M. Wolf, PhD; Elaine M. Hockman, PhD; Nancy P. Allen, PhD; C. Victor Bunderson, PhD; Russell T. Osgethorpe, PhD; and Sally H. Cavanaugh, PhD for their insightful review and recommendations to improve this work; and to Becky Bluett and Christine Downs for their assistance in preparing the manuscript.

Funding: In part by a University of Michigan curriculum innovation grant.

Conflicts of Interest: No personal, commercial, political, academic or financial conflict of interest was reported by any of the authors. DMC is editor-in-chief of JCRPC.

Ethical Approval: Authors report the study qualified for IRB exemption.

Correspondence: Dane M. Chapman, MD, PhD, Director, Center for Clinical Reasoning and Procedural Competency, Department of Emergency Medicine, University of Missouri-Columbia School of Medicine, 1 Hospital Drive, Columbia MO, 65201, e-mail: <chapmandan@health.missouri.edu>

References

1. AAMC Report of the Ad Hoc Committee of Deans. Educating doctors to provide high quality medical care. 2004: 1-12. Available from: https://members.aamc.org/eweb/upload/Educating%20Doctors%20to%20Provide%20July%202004.pdf.
2. Epstein RM, Hundert EM. Defining and assessing professional competence. *JAMA*. 2002;287(2):226-35.
3. Chapman DM, Calhoun JG, Van Mondfrans AP, Davis WK. Measuring clinical reasoning competency using a virtual patient model. *JCRPC*. 2013;1(1):1-9
4. Chapman DM, Char DM, Aubin C. Clinical decision making. Chapter 10 in: Marx JA et al., editors. *Rosen's Emergency Medicine, Concepts and Clinical Practice* 6th Edition. Philadelphia, PA: Mosby/Elsevier; 2006:125-33.
5. Norman GR. The epistemology of clinical reasoning: Perspectives from philosophy, psychology, and neuroscience. *Acad Med.* 2000;75(10):S127-33.
6. Elstein AS. Clinical problem solving and decision psychology: Comment on "The epistemology of clinical reasoning." *Acad Med*. 2002;75(10):S134-36.
7. Elstein AS, Schwarz A. Clinical problem solving and diagnostic decision making: Selective review of the cognitive literature. *BMJ* 2002;324:729-32.
8. Hoffman K, Hosokawa M, Blake R, Headrick L, Johnson G. Problem-based learning outcomes: Ten years of experience at the University of Missouri-Columbia School of Medicine. *Acad Med*. 2006;81:617-625.
9. Patel VL, Arocha JF, Leccisi, MS. Impact of undergraduate medical training on housestaff problem-solving performance: Implications for problem-based curricula. *J Dent Educ.* 2001;65(11):1199-1218.
10. Friedman CP. The marvelous medical education machine or how medical education can be unstuck in time. *Acad Med*. 2000;75(10):S137-42.
11. Stevens SM, Goldsmith TE, Summers KL et al. Virtual reality training improves students' knowledge structures of medical concepts. In: Westwood JD et al., editors. *Medicine Meets Virtual Reality* 13. Amsterdam: IOS Press; 2005:519-25.
12. Jarrell BE. Simulation for teaching decision making in medicine: The next step. Presented at Medicine Meets Virtual Reality 13, Long Beach, CA, Jan 28, 2005.
13. Cook DA, Triola MM. Virtual patients: A critical literature review and proposed next steps. *Med Educ*. 2009;43(4):303-11.
14. BMDP Statistical Software, Inc. 1964 Westwood Blvd, Suit 202, Los Angeles, CA. Available from: http://www.statistical-solutions-software.com/bmdp-statistical-software/bmdp/.
15. Gruppen LD, Palchik NS, Wolf FM, et al. Medical student use of history and physical information in diagnostic reasoning. *Arthritis Care Res*. 1993;6(2):64-70.
16. Chapman DM, Hayden S, Sanders AB, Binder LS, Chinnis A, Corrigan K, LaDuca T, Dyne P, Perina DG, Smith-Coggins R, Sulton L, Swing S. Integrating the Accreditation Council for Graduate Medical Education core competencies into the Model of the Clinical Practice of Emergency Medicine. *Ann Emerg Med.* 2004;33(6):756-769.
17. Eshach H, Bitterman H. From case-based reasoning to problem-based learning. *Acad Med.* 2003;78(5):491-96.
18. Tärnvik A. Revival of the case method: A way to retain student-centred learning in a post-PBL era. *Med Teach*. 2007;29:32–36.
19. Chapman DM. Teaching and evaluating clinical reasoning through computer-based patient management simulations. *Dissertation Abstracts International.* 1985;46:784-B.
20. Gesundheit N, Brutlag P, Youngblood P, et al. The use of virtual patients to assess the clinical skills and reasoning of medical students: Initial insights on student acceptance. *Med Teach.* 2009;31(8): 739-42.

Assessing Effectiveness of a Problem-Based Learning Curriculum in Teaching Clinical Reasoning Skills

DANE M. CHAPMAN, MD, PhD[1,2,3]
JUDITH G. CALHOUN, PhD, MBA[1]
ADRIAN P. VAN MONDFRANS, PhD[2]
WAYNE K. DAVIS, PhD[1]

Institutional Affiliations:

1. *University of Michigan Medical School, Department of Medical Education, Ann Arbor, MI*
2. *Brigham Young University, McKay School of Education, Department of Instructional Psychology and Technology, Provo, UT*
3. *University of Missouri-Columbia, School of Medicine, Center for Clinical Reasoning and Procedural Competency, Department of Emergency Medicine, Columbia, MO*

Abstract

Background:
Problem-based learning has been advocated in teaching clinical reasoning, yet it is unclear how to best measure clinical reasoning skills using this approach.

Objective:
To evaluate the efficacy of free-inquiry, virtual patient simulations compared to menu-driven, branching, written patient simulations in assessing data-gathering and clinical decision-making skills during a four-week, post-second year problem-based learning curriculum.

Methods:
Experimental, multiple-groups pretest-posttest control group and quasi-experimental pretest-posttest control group designs were used to evaluate expected improvements in clinical reasoning over the four-week curriculum. All post-second year medical students (N=191) were required to participate in the problem-based learning curriculum and complete a written patient simulation at the end of each week. Ninety-seven of the 191 medical students volunteered without compensation to participate in an additional 6.5-hour virtual patient simulation elective during the problem-based curriculum and were randomly assigned to one of the four weeks. Simulation elective students completed three virtual patient practice simulations with feedback and three assessment simulations (one pretest and two posttests) measuring nineteen dependent variables from four clinical reasoning constructs: proficiency, efficiency, thoroughness and strategy.

Results:
Multivariate, repeated-measures, factorial ANOVA statistics revealed a significant problem-based learning curriculum effect upon the second virtual patient simulation posttest along all four clinical reasoning constructs: proficiency ($p = .03$), efficiency ($p = .01$), thoroughness ($p = .00$), and strategy ($p = .01$). For three of the four constructs (proficiency, efficiency and strategy), no significant differences among the four weeks were found on multivariate analyses of the virtual patient pretest, suggesting a combined problem-based curriculum and virtual patient simulation practice enhancement of clinical reasoning competency. These enhancements were not detected by the written patient simulations.

Conclusions:
A four-week, problem-based learning curriculum can significantly enhance clinical reasoning competency and three hours of virtual patient simulation practice augments that effect. Results also support the utility of free-inquiry, virtual patient simulations in teaching and assessing clinical reasoning competency.

Keywords:
problem-based learning, clinical reasoning, competency assessment, virtual patient simulations, medical students

Introduction

In response to public pressure for greater accountability from the medical profession, medical schools and residency programs are undergoing major transformations to ensure students and residents learn what they need to know to become competent physicians.[1,2] Coupled with this accountability response, undergraduate and graduate medical education programs and certifying boards have stepped up efforts to assure physician competency through outcomes measurement.[3-4] While few would argue the importance of an outcomes approach to competency assessment, uncertainty exists on how to best teach and evaluate clinical reasoning skills of data gathering (i.e. history taking, physical examinations and selecting and interpreting diagnostic tests) and diagnostic and therapeutic decision making.[3-5]

Problem-based learning has been advocated for teaching clinical reasoning skills. Initially described by Neufeld and Barrows,[6] problem-based learning (PBL) has been implemented to varying degrees within the more conventional medical school curriculum. Of the eighty percent of U.S. medical schools that report using PBL, 45 percent report fewer than 10 percent PBL preclinical contact hours.[7] In a meta-analysis of 20 years of PBL experience (1972-1992) compared with conventional students, PBL graduates demonstrated gaps in their cognitive knowledge base and used a more backward rather than forward clinical reasoning approach characteristic of expert clinicians.[8] More recently, Hoffman et al.[9] using an outcomes approach, describe their ten year experience using PBL at the University of Missouri-Columbia School of Medicine from 1993-2006. They report significantly higher USLME Step 1 and Step 2 scores for their PBL students compared to first-time examinees nationally. These gains in performance appear to continue into residency as program directors note superior performance of the school's PBL graduates.[9] However, they do not report how well PBL graduates did on the USMLE Step 3 Computer-based Case Simulation (CCS) exam designed to measure an examinees' approach to clinical management, including diagnosis, treatment and monitoring.[9]

Uncertainty remains regarding the best method for measuring PBL effectiveness in teaching clinical reasoning.[3-5] Based upon a review of the past thirty-five years of professional medical assessment, new, more reliable modalities are recommended for assessing clinical reasoning and expert judgment among other professional attributes.[10] Fifteen years ago, the Accreditation Council for Graduate Medical Education (ACGME) initiated the "ACGME Outcome Project" to shift the focus of residency program requirements and accreditation from process-oriented assessment to an assessment of educational outcomes of resident and residency program performance as a basis for accreditation.[11] A tidal wave of change has come as a result of that initiative including the ACGME core competencies,[12] the ACGME toolbox,[13] and controversy over the psychometric inadequacy of the tool box instruments[13] and whether the core competencies themselves represent valid measurement constructs.[14]

Fueled by this massive push in educational outcomes assessment and attendant controversy, a critical need exists for psychometrically proven assessment modalities. Despite increasing popularity of computer-based or virtual patient simulations in medical schools and residencies, their role in teaching and assessing clinical skills and diagnostic and therapeutic decision making requires greater clarificaiton.[10,15-18] To maximize instructional validity (i.e. that the test actually measures what it purports to measure), an optimal virtual patient simulation will be high fidelity—meaning it will faithfully simulate the actual physician-patient encounter. It will also be free-inquiry—meaning users can access data freely without menus or other branching limitations, and without cues.[16-18] Rather than text or verbal descriptions of physical exam and diagnostic test findings, actual visual and auditory responses will be provided, such as visual cues for skin rashes, cardiac and respiratory sounds, and digital images for electrocardiographs (EKG's) and radiographs.

This study aims to evaluate whether: (1) high-fidelity, free-inquiry virtual patient (VP) simulations compared to menu-driven, branching written patient (WP) simulations are better in detecting the impact of a four week PBL curriculum (PBLC) upon clinical reasoning (CR) skill acquisition of post-second year medical students; and (2) three hours of VP simulation practice with feedback would further impact CR skill acquisition beyond the PBL curriculum alone. It was hypothesized that VP compared to WP simulations would be more sensitive (instructionally valid) in detecting significant improvements in clinical reasoning competency over the four-week PBLC, and that VP simulation practice would increase CR skill acquisition over PBLC alone.

Methods

Study Design:

We used an experimental, multiple-groups, pretest-posttest control group design[19] to assess the effect of the required four-week PBL curriculum and three hours of VP simulation practice (independent variables) upon nineteen previously established clinical reasoning skill indices[20] (dependent variables). The multiple-groups experimental pretest-posttest design was chosen to control for the influence of historical events and non-PBLC related maturation throughout the four weeks of the PBLC. Random assignment of students to groups made it unlikely that a consistent effect throughout the four weeks of the PBLC would be due to historical or maturational events unrelated to the PBL curriculum itself. As an additional control, a second analysis using a quasi-experimental pretest-posttest control group design[19] compared PBLC-only students with those students who also experienced VP simulation practice and assessment in addition to the PBL curriculum.

Study Setting and Population:

The study qualified for institutional review board (IRB) exemption as a curriculum innovation project and was conducted at the University of Michigan Medical School. The entire medical school class of post-second year medical students (N=191) was required to take the PBL curriculum between the preclinical and clinical years of medical school. Ninety-seven of the 191 post-second-year medical students volunteered without compensation to participate in a virtual patient simulation elective during the required four-week PBL curriculum and formed the VP elective group. The remaining 94 students formed the PBLC-only control group. The entire class of 191 students underwent random assignment to each week of the four week PBLC stratified on VP elective participation, gender and minority status such that of the 46-47 students randomly assigned to each week of the PBLC, 23-25 were randomly assigned to the VP elective.

Problem-based Learning Curriculum:

The PBL curriculum was a required, four-week block from May 4th to June 1st between the second and third years of medical school. It was "designed to provide a relevant practical clinical problem solving experience for students about to enter the clinical phase, with emphasis on the philosophy of lifetime learning and independent and group thought in the practice of clinical problem solving and diagnosis."[21] It consisted of expert presentations on the practice and concepts of clinical problem solving and physical diagnosis with faculty and students participating in both lecture and small group discussions. The PBLC sequence was originally developed around eight organ-system themes with each of the four weeks of the PBLC focusing upon two major organ systems: Week 1, cardiovascular and pulmonary; Week 2, gastrointestinal and renal; Week 3, neurologic and musculoskeletal; and Week 4, endocrine and reproductive.[22] The objectives of the PBLC were to: (1) bridge the gap between the basic science and clinical practice years of medical school, (2) review patient cases and provide practice in clinical problem solving and diagnosis, (3) emphasize clinical assessment skills, (4) provide mini-elective opportunities such as the VP elective, and (5) introduce medical students to the hospital setting during a clinical week.

During the mornings of their clinical week, PBLC students saw patients and functioned as part of the hospital ward team. All 191 students were expected to attend regularly scheduled afternoon dialectic sessions on Monday, Wednesday and Friday afternoons from 1:00 to 4:00 PM throughout the four-week PBLC. Completion of a weekly WP simulation was required of all PBLC students during the Friday afternoon dialectic session. VP students experienced their VP elective week during a week other than their clinical week. The 23-25 VP students randomly assigned to each week of the PBLC underwent further random assignment into one of four treatment subgroups.

Virtual Patient Simulation Elective:

The 6.5 hour VP elective included two sessions (3.0 and 3.5 hours) on Monday through Wednesday, Tuesday through Thursday or Wednesday through Friday mornings, during which students worked through six VP simulations: one 60-minute pretest (Case 1: Cardiology), three 60-minute practice simulations with corrective feedback (Case 2: Pediatric Endocrinology, Case 3: Infectious Disease and Case 4: Pulmonary), and two 45-minute posttests (Case 5: Pulmonary and Case 6: Cardiology). No corrective feedback was provided for pretest or posttest assessment simulations. VP students were randomly assigned to work individually or in groups of three during practice simulations. All VP students completed their pretest and posttest simulations as individuals.

Written Patient Simulations:

A different WP simulation was administered each week to all PBLC students during the Friday afternoon dialectic session. The WP simulations were developed by the University of Michigan Medical School following a menu-driven and branching format used by the National Board of Medical Examiners.[23] Each WP simulation began with an "opening scene" which introduced the patient and presented the chief complaint, the setting and the students' role. Each opening scene was followed by five sections representing steps in the workup, diagnosis and management of the patient. These sections were: A) History, B) Physical Examination, C) Diagnostic Studies, D) Differential and Principal Diagnoses, and E) Therapeutic Procedures. The first three sections incorporated 15-40 decision options. Some options were appropriate or indicated, others were inappropriate or contraindicated, and still others were optional or neutral.

Virtual Patient Simulations:

The multi-problem, network-based VP simulations used in this study simulated the actual physician-patient encounter with high fidelity and free inquiry, without menu-cueing or branching limitations. They were derived from content experts, had documented critical actions for data gathering and diagnostic elements, and could be altered to prevent corrective feedback during the assessment cases. Following an "opening scene," users assumed the role of physicians and moved to and from history, physical examination, diagnostic study, diagnosis and treatment sections without menu-driven cueing or branching limitations.[20]

The VP simulations were not the ultimate virtual patient, however, as artificial intelligent responses to all history, physical exam and diagnostic test inquiries were provided as text, and not virtual touch, sound, or images.

Assessments and Procedure:

WP Simulation: An individual score for each WP simulation was represented by a percentage of the correct points for each of the five sections. A "problem-solving index" was used as an overall performance score on each WP simulation and consisted of the average of section scores across the five sections.[22-23] Concurrent, criterion-referenced and instructional validities for the WP simulations have been previously reported.[22] The WP simulation assessments were administered on the Friday of each PBLC week simultaneously to VP elective and PBLC-only groups in a standard lecture hall of the University of Michigan Medical School.

VP Simulation: To standardize transcript scoring, coding regulations were developed using sample VP simulation transcripts. Case-specific VP simulation scoring protocols provided a summary of those expert-recommended critical inquiries that had been made. Diagnosis sections were independently scored by two individuals using case-specific coding regulations that identified acceptable synonyms for diagnoses. Transcripts were scored by at least one rater who was blinded to pretest-posttest classification, and inter-rater agreement was consistently high (r>.90). While therapeutic and management plans were also computer-scored, these were ignored for the purposes of this study.

The VP simulation pretest (Case 1: Cardiology) and both posttests (Case 5: Pulmonary and Case 6: Cardiology) were used to measure the effect of the four-week PBL curriculum upon the nineteen clinical reasoning competency indices, while controlling for any VP simulation pretest-posttest differences due

TABLE 1: Mathematical Descriptions of Nineteen Clinical Reasoning Performance Indices Derived for Use in Multi-Problem Virtual Patient Simulations

Index	Abbreviation	Description[a]
Proficiency		
History Taking	HTP	(Obtained CHT/Total CHT) X 100
Physical Examination	PEP	(Obtained CPE/Total CPE) X 100
Diagnostic Tests	DTP	(Obtained CDT/Total CDT) X 100
Correct Diagnoses	DP	(Obtained CD/Total CD) X 100
Program Solving Index	PSI	(HTP + PEP + DTP + DP) / 4
Proficiency Index	PI	(Obtained CHT + CPE + CDT) X 100 / (Total CHT + CPE + CDT)
Efficiency		
History Taking	HTE	(CHT Obtained/HTT) X 100
Physical Examination	PEE	(CPE Obtained/PET) X 100
Diagnostic Tests	DTE	(CDT Obtained/DTT) X 100
Thoroughness		
History Taking	HTT	Total HT
Physical Examination	PET	Total PE
Diagnostic Tests	DTT	Total DT
Total Data-Gathering	TDG	(HTT + PET + DTT)
Diagnosis	DT	Total D
Strategy		
History Taking	HTS	[HTT/(HTT+PET + DTT)] X 100
Physical Examination	PES	[PET/(HTT+PET + DTT)] X 100
Diagnostic Tests	DTS	[DTT/(HTT+PET + DTT)] X 100
Focused Strategy Index	FSI	(HH + PP+ DD +1) / (HP + HD + PH + PD + DH + DP + 1)
Invasiveness/Cost Index	ICI	[DTT/(HTT + PET)] X 100

[a]*Symbol Key: HT*= history taking inquiries, *PE*= physical examination inquiries, *DT*= diagnostic test inquiries, *D*= diagnoses indicated, *C*= critical inquiry or diagnosis (e.g. CHT=critical history taking inquiries), *HH*= history to history transition, *PP*= physical exam to physical exam transition, *DD*= diagnostic test to diagnostic test, *HP*= history to physical exam, HD= history to diagnostic test, *PH*= physical exam to history, *PD*= physical exam to diagnostic test, *DH*= diagnostic test to history, and *DP*= diagnostic test to physical exam transition.

to the instructional effect of the three VP practice simulations. The VP simulation pretest was administered during the first VP Session on Monday, Tuesday or Wednesday of each of the four PBLC weeks. The two VP simulation posttests were administered during the second VP session on Wednesday, Thursday or Friday. VP simulation assessments were administered in the Taubman Health Sciences Library Learning Resource Center. Proctors were available to respond to VP simulation computer interface or program questions.

VP Simulation Clinical Reasoning Indices: CR performance scores along nineteen predetermined CR indices[20] (dependent variables) were used to assess the effect of PBLC and VP simulation practice over the four weeks of the PBLC. The nineteen CR competency indices were previously validated and are represented as mathematical descriptions under their corresponding clinical reasoning construct: proficiency, efficiency, thoroughness, and strategy (See Table 1).

Clinical reasoning proficiency referred to how effectively critical data were gathered and correct diagnoses made. The CR proficiency indices were: percent of critical data-gathering inquiries obtained for history (history proficiency), physical examination (physical examination proficiency), and diagnostic tests (diagnostic test proficiency); percent of correct diagnoses made (diagnosis proficiency); Problem-Solving Index (PSI)—an average of data-gathering and decision-making proficiencies; and Proficiency Index (PI)—the percent of data-gathering critical information obtained.

Clinical reasoning efficiency was defined as the percentage of data-gathering inquiries that were critical in making the diagnosis of a patient's problem(s). Higher scores represented greater efficiency in making medical inquiries. Clinical reasoning efficiency indices included history, physical examination and diagnostic test efficiencies.

Clinical reasoning thoroughness reflected the frequency of data-gathering inquiries made or diagnoses indicated. Clinical reasoning thoroughness indices included: total number of history inquiries (history thoroughness), physical examination inquiries (physical examination thoroughness), and diagnostic test inquiries (diagnostic test thoroughness); total number of history, physical examination and diagnostic test inquiries combined (total data-gathering thoroughness); and total number of diagnoses hypothesized at the completion of each simulated case (diagnosis thoroughness).

Clinical reasoning strategy referred to the cognitive strategies used to arrive at correct diagnoses. It reflected individual preference for certain data-gathering techniques (e.g. to use either a focused inquiry approach or a "shot-gun" or haphazard approach). CR strategy indices included: percent of total data-gathering inquiries that relate to history taking (history strategy), physical examination (physical examination strategy), or diagnostic test (diagnostic test strategy); Focused Strategy Index—the standardized proportion of data-gathering inquiry transitions of similar type (e.g. history to history) to all other combinations of possible inquiry transitions from one type of inquiry to another (e.g. history to physical examination, diagnostic test to history, etc.), where high scores reflect a more focused and systematic data-gathering approach; and Invasiveness/Cost Index—the standardized proportion of diagnostic test inquiries (relatively invasive and costly) to the sum of history-taking and physical examination inquiries (relatively non-invasive and less costly), where higher scores reflect a more invasive and costly data-gathering approach.

Data Analysis:

BMDP multivariate, repeated measures, factorial ANOVA statistics were used to determine any pretest-posttest differences in CR constructs due to VP simulation practice while controlling for week of the PBL curricululm (independent and control variables). If a significant multivariate effect was observed, then univariate ANOVA statistics were used to test for differences along the nineteen clinical reasoning dependent variables.[20] Univariate, factorial ANCOVA statistics were used to compare VP elective and PBLC-only groups over PBL curriculm weeks 1 through 4. (http://www.statistical-solutions-software.com/bmdp-statistical-software/bmdp/).

Results

Complete data for 184 of the 191 PBL curriculum students were available for comparing VP elective and PBLC-only groups over the four weeks of the PBL curriculum. The VP elective students (N=97) appeared to be representative of the entire medical school class (N=191) as VP students did not differ significantly from other class members in ethnicity, sex, or prior clinical experience on the hospital wards. Approximately two-thirds of the VP enrollees had never previously participated in computer-based instruction and almost one-fourth had never interacted with a computer in any capacity at the time of the original study, making a selection bias, which favored students who were more comfortable with using computers for learning, unlikely.

Computer transcripts (N=486) were generated by 81 VP medical students with complete data, and documented student-computer interactions for 243 hours of practice and 202 hours of assessment. Outcome performance scores along nineteen predetermined CR indices[20] (dependent variables) were derived from the 243 hard-copy VP simulation assessment transcripts (one pretest: Case 1; and two posttests: Case 5 and Case 6) and are represented as means and standard deviations (SD) (See Table 2).

Detecting Problem-Based Learning Curriculum Impact

TABLE 2: Pretest-Posttest Means (SD) for Nineteen Clinical Reasoning Indices Over Four Weeks of the Problem-Based Learning Curriculum (N=81)[a]

Index	Week	Case 1 Pretest (Cardiology)	Case 5 Posttest (Pulmonary)	Case 6 Posttest (Cardiology)
Proficiency				
History Taking	1	38.1 (24.5)	57.7 (17.4)[c]	44.3 (15.0)
	2	50.0 (26.2)	55.9 (19.6)	52.8 (13.8)
	3	48.5 (25.7)	60.3 (19.4)	51.2 (14.5)
	4	44.3 (23.0)	51.7 (20.1)	48.2 (16.2)
Physical Examination	1	46.4 (27.2)	65.1 (19.7)[c]	52.4 (20.5)[cd]
	2	54.8 (17.0)	60.3 (13.4)	62.8 (19.3)
	3	57.3 (24.6)	72.6 (17.6)	62.3 (13.9)
	4	46.6 (20.8)	71.2 (18.7)	70.9 (19.2)
Diagnostic Tests	1	47.6 (23.6)	49.0 (10.6)[b]	42.3 (13.9)[bd]
	2	58.3 (31.0)	55.1 (17.0)	56.6 (13.6)
	3	69.1 (18.8	52.1 (7.0)	54.2 (20.4)
	4	61.4 (25.3)	50.6 (11.4)	57.1 (15.5)
Correct Diagnoses	1	28.6 (26.4)	55.9 (22.2)[c]	30.9 (20.8)[b]
	2	39.7 (27.1)	44.0 (23.6)	35.7 (18.7)
	3	29.4 (20.0)	47.0 (17.4)	26.5 (10.7)
	4	48.5 (26.7)	40.9 (19.7)	27.3 (17.1)
Program Solving Index	1	40.2 (13.1)	56.9 (8.6)[c]	42.5 (9.1)[e]
	2	50.7 (17.0)	53.8 (11.6)	52.0 (9.8)
	3	51.1 (12.0)	58.0 (10.9)	48.6 (8.1)
	4	50.2 (14.8)	53.6 (10.4)	50.9 (9.1)
Proficiency Index	1	44.0 (14.5)	55.6 (8.9)[b]	45.2 (7.8)[e]
	2	54.4 (18.0)	56.4 (12.5)	56.3 (9.2)
	3	58.3 (12.8)	59.1 (10.0)	54.6 (10.1)
	4	50.7 (16.0)	54.5 (11.5)	56.2 (9.3)
Efficiency				
History Taking	1	7.3 (4.4)	20.4 (8.7)[c]	26.3 (12.2)[c]
	2	8.8 (5.3)	20.2 (9.8)	22.7 (7.6)
	3	8.2 (4.0)	19.1 (7.8)	26.7 (7.5)
	4	8.4 (5.9)	17.8 (8.4)	24.9 (14.2)
Physical Examination	1	18.5 (9.2)	18.8 (6.7)	25.3 (10.9)[cd]
	2	21.3 (10.6)	19.6 (11.5)	24.7 (12.1)
	3	15.3 (5.9)	14.5 (8.3)	17.2 (8.4)
	4	15.4 (6.2)	16.3 (12.7)	19.8 (6.5)
Diagnostic Tests	1	40.4 (23.8)	45.1 (18.9)[c]	33.8 (9.6)
	2	32.3 (25.2)	38.8 (16.7)	32.2 (12.4)
	3	34.3 (14.2)	46.1 (13.0)	40.0 (18.2)
	4	34.5 (25.0)	40.3 (14.6)	39.0 (13.1)

[a]Repeated-Measures multivariate factorial ANOVA for Pretest - Posttest Comparisons, controlling for significant PBLC Week Effect for 81 medical students with complete data, significant pretest-posttest effect; [b]$p \leq 0.050$; [c]$p \leq 0.010$; significant PBLC Week effect [d]$p \leq 0.050$; [e]$p \leq 0.010$; [f]standardized T-Score mean=50, SD=10

Index	Week	Case 1 Pretest (Cardiology)	Case 5 Posttest (Pulmonary)	Case 6 Posttest (Cardiology)
Thoroughness				
History Taking	1	21.9 (8.5)	25.4 (11.2)[b]	20.8 (11.4)
	2	24.5 (10.8)	25.8 (11.0)	25.4 (8.8)
	3	23.3 (9.2)	27.8 (11.8)	20.3 (7.3)
	4	24.9 (10.5)	26.0 (10.6)	22.6 (11.7)
Physical Examination	1	10.0 (5.0)[d]	12.1 (5.2)[cd]	11.7 (5.5)[ce]
	2	11.5 (4.7)	12.8 (7.0)	14.9 (6.8)
	3	15.2 (5.0)	18.1 (6.7)	18.4 (3.4)
	4	12.8 (6.2)	17.4 (8.9)	18.9 (5.8)
Diagnostic Tests	1	5.5 (2.6)[e]	8.2 (3.0)[b]	11.9 (4.2)[cd]
	2	9.0 (4.4)	11.0 (4.8)	18.2 (8.0)
	3	9.8 (5.3)	8.6 (2.8)	13.6 (5.9)
	4	9.7 (4.7)	9.7 (3.3)	14.2 (5.6)
Total Data-Gathering	1	37.4 (11.1)[d]	45.7 (14.1)[c]	44.3 (12.4)[cd]
	2	45.0 (13.2)	49.6 (17.3)	58.5 (14.9)
	3	48.2 (11.7)	54.5 (15.4)	52.3 (9.5)
	4	47.4 (12.8)	53.0 (17.4)	55.8 (15.1)
Diagnosis	1	2.5 (1.1)[e]	3.0 (0.8)	4.3 (1.2)[c]
	2	3.7 (1.3)	3.3 (1.5)	5.8 (2.5)
	3	2.8 (0.8)	3.4 (1.1)	4.7 (1.5)
	4	3.1 (1.2)	2.6 (0.9)	4.5 (1.6)
Strategy				
History Taking	1	58.1 (11.3)	54.2 (11.3)	44.7 (16.3)[c]
	2	52.8 (12.8)	51.6 (8.6)	43.7 (12.9)
	3	47.5 (10.8)	50.3 (10.5)	38.3 (10.2)
	4	51.3 (12.8)	48.1 (8.9)	38.2 (13.3)
Physical Examination	1	26.4 (8.6)	26.6 (7.9)[e]	26.6 (10.6)[be]
	2	26.2 (9.6)	24.8 (6.7)	25.1 (8.6)
	3	31.4 (7.1)	33.0 (8.4)	36.0 (8.2)
	4	26.6 (9.2)	31.3 (9.0)	34.2 (6.6)
Diagnostic Tests	1	15.5 (7.1)	19.3 (8.0)[d]	28.7 (12.0)[c]
	2	20.9 (10.1)	23.6 (9.4)	31.2 (11.9)
	3	21.0 (11.5)	16.7 (6.0)	25.7 (8.4)
	4	22.0 (13.8)	20.5 (9.5)	27.6 (13.4)
Focused Strategy Index[f]	1	49.0 (8.4)	50.8 (9.3)[c]	47.7 (8.7)[c]
	2	49.6 (8.6)	47.3 (10.4)	50.3 (12.5)
	3	51.9 (11.7)	51.4 (7.9)	50.3 (7.8)
	4	52.6 (12.7)	50.7 (12.2)	51.0 (10.1)
Invasiveness/Cost Index[f]	1	45.8 (4.7)	48.5 (7.5)[d]	50.3 (9.7)[c]
	2	50.0 (7.7)	52.8 (9.1)	52.1 (10.3)
	3	50.5 (9.4)	45.9 (4.9)	47.6 (6.0)
	4	52.2 (13.4)	50.0 (9.7)	50.2 (12.5)

[a]Repeated-Measures multivariate factorial ANOVA for Pretest - Posttest Comparisons, controlling for significant PBLC Week Effect for 81 medical students with complete data, significant pretest-posttest effect; [b]$p \leq 0.050$; [c]$p \leq 0.010$; significant PBLC Week effect [d]$p \leq 0.050$; [e]$p \leq 0.010$; [f]standardized T-Score mean=50, SD=10

Using VP Simulation Assessments:

Multivariate, repeated measures, factorial ANOVA revealed a significant week of PBLC effect for the second posttest (Case 6: Cardiology) across all four clinical reasoning constructs: proficiency (Wilks' λ=.61, F(18,170)=1.80; p = .03), efficiency (Wilks' λ=.73, F(9,153)=2.38; p = .01), thoroughness (Wilks' λ=.54, F(12,164)=3.53; p = .00), and strategy measures (Wilks' λ=.73, F(15,168)=2.38; p = .01) (See Table 2). A significant week of PBLC effect was also noted for the first posttest (Case 5: Pulmonary) along thoroughness (Wilks' λ=.66, F(12,164)=2.36; p = .00), and strategy constructs (Wilks' λ=.61, F(15,168)=2.21; p = .01); and for the pretest (Case 1: Cardiology) along the thoroughness construct (Wilks' λ=.59, F(12,164)=2.99; p = .00) (See Table 2). Since clinical reasoning competency differences over the weeks of the PBLC were noted for each of the four clinical reasoning constructs on the second posttest only, follow-up univariate ANOVAs were performed on the second posttest only.

Univariate ANOVA resulted in significantly higher proficiency posttest scores among the four groups (PBLC week 1 through 4) in physical exam (F(3,65)=3.25, p = .027), diagnostic tests (F(3,65)=3.96, p = .012), Problem-Solving Index (F(3,65)=4.56, p = .006) and Proficiency Index (F(3,65)=6.66, p = .000), but not in history taking or diagnostic accuracy proficiency indices (see Table 2, Figure 1). Univariate ANOVA of efficiency scores detected a significant decrease rather than increase in physical examination efficiency over the four-week PBLC (F(3,65)=3.51, p = .020), suggesting that even though students were making more overall physical examination inquiries, they were not obtaining proportionally more critical physical examination findings (See Figure 1). Univariate ANOVA demonstrated significantly higher posttest thoroughness scores over the four weeks of the PBLC in physical examination (F(3,65)=6.65, p = .000), diagnostic study (F(3,65)=3.45, p = .021), and total data gathering thoroughness (F(3,65)=3.69, p = .016), but not in history thoroughness (See Figure 2). Univariate ANOVA of strategy measures revealed a tendency to use proportionally more physical examination inquiries compared to history taking or diagnostic study inquiries over the four-week PBLC (F(3,65)=7.30, p = .000) (See Figure 2).

Detecting Problem-Based Learning Curriculum Using WP

FIGURE 1: Clinical Reasoning Proficiency and Efficiency of Preclinical Medical Students (N=81) Using a Virtual Patient Simulation Over Four Weeks of the PBLC (*ANOVA, F Test)

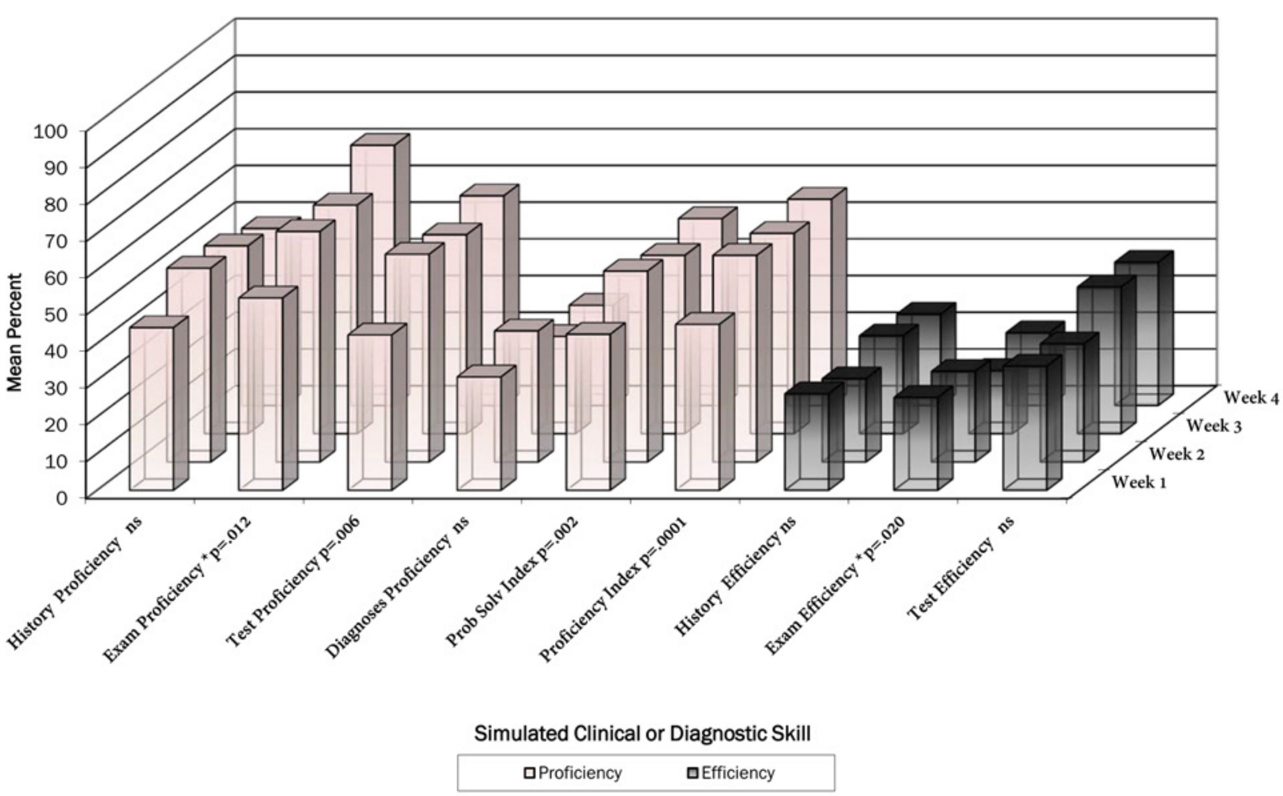
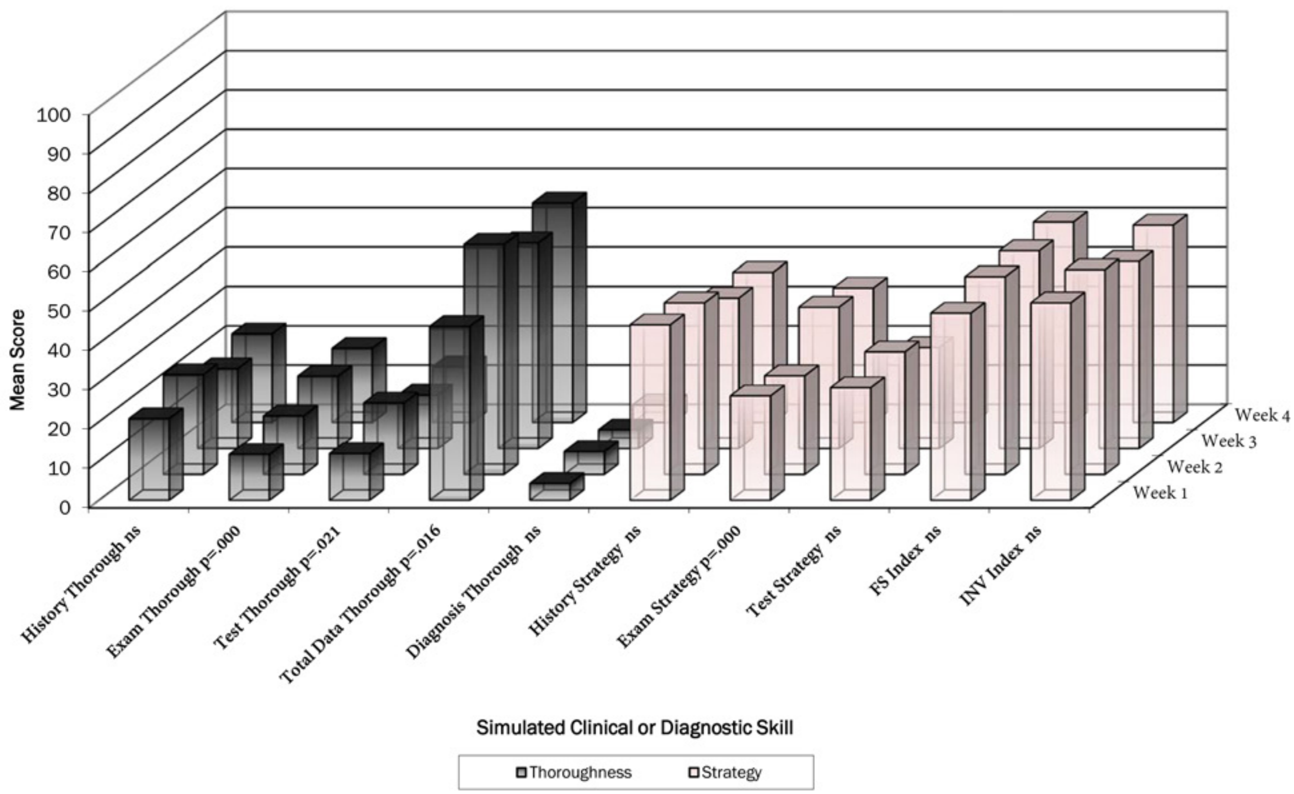


FIGURE 2: Clinical Reasoning Thoroughness and Strategy of Medical Students (N=81) Using a Virtual Patient Simulation Over Four Weeks of the PBLC (*ANOVA, F Test)



Simulation Assessments:

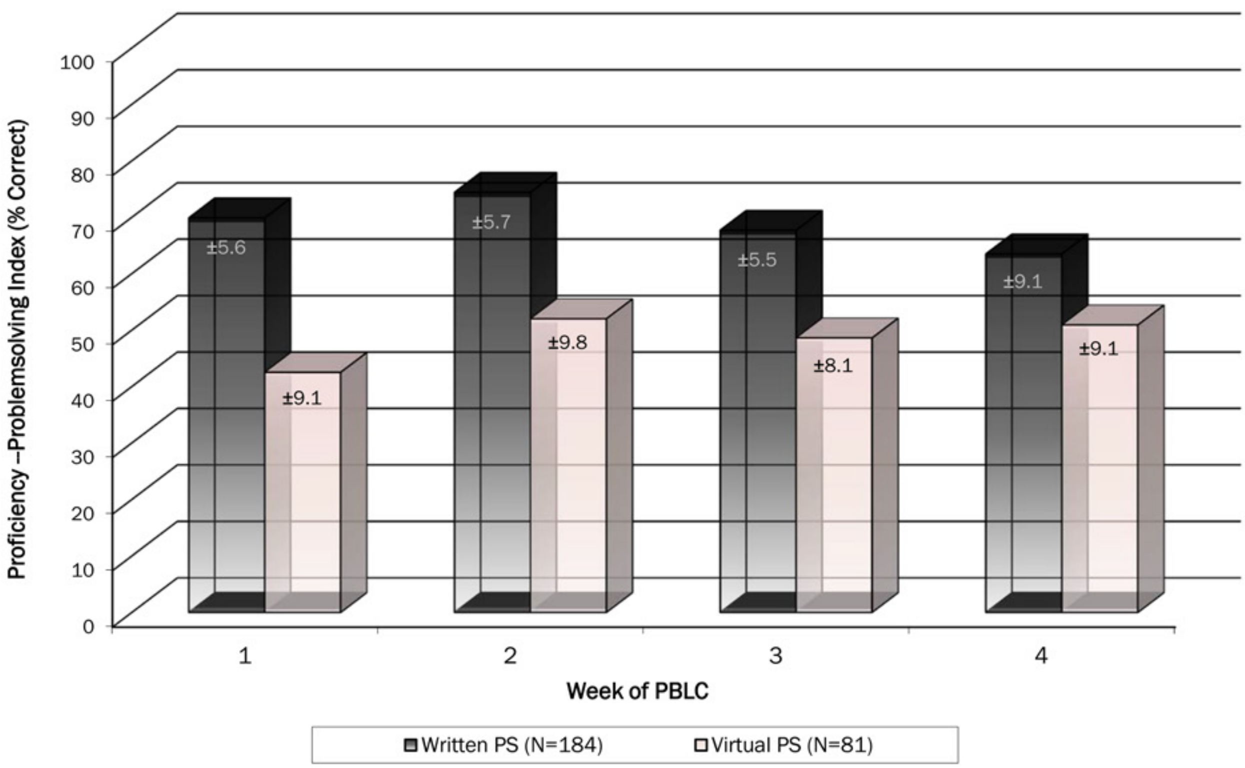
A strong PBL curriculum week effect was noted on all six WP simulation proficiency measures including History Taking ($F(3,735)=55.32$; $p=.00$), Physical Examination ($F(3,735)=147.9$; $p = .00$), Diagnostic Examination ($F(3,735)=46.78$; $p = .00$), Principal and Differential Diagnoses ($F(3,735)=70.33$; $p = .00$), Therapeutic Procedures ($F(3,735)=138.9$; $p = .00$), and Problem-Solving Index ($F(3,735)=99.83$; $p = .00$). However, rather than demonstrating improved clinical reasoning performance over each of the four weeks of the PBL curriculum, medical student (N=184) proficiency scores tended to decrease over the four weeks of the PBL curriculum (See Figure 3); suggesting WP simulation assessments were of varying difficulty level and were likely confounding our results.

Using multivariate ANOVA statistics, no differences in VP elective and PBLC-only performance were found for any of the six WP simulation proficiency measures over any of the four weeks of the PBL curriculum. To better control for any confounding pretreatment differences among students who enrolled in the VP elective and those who did not, and to increase the precision of comparisons between the VP elective and PBLC-only groups, a repeat analysis was done using univariate factorial analysis of covariance (ANOVA) statistics in a pretest-posttest control group design with the pretest (Week 1 WP simulation) being treated as the covariate. Comparisons were made between VP elective and PBLC-only groups for weeks 2, 3 and 4 of the PBL curriculum. Using this approach, no significant differences were noted between the VP elective and PBLC-only groups for any of the six WP simulation proficiency measures for any week of the PBL curriculum ($p>.05$, ANCOVA).

Discussion

The fact that both WP and VP simulation scores for CR constructs varied significantly over the four weeks of the PBLC supports the PBL curriculum impact on CR competency. Since no differences were found between VP elective and PBLC-only students, groups were likely equal in clinical reasoning proficiency during each week of the PBLC. If three hours of VP simulation practice was truly enhancing CR competency as measured by VP simulation assessments (See Table 2), such improvements were not being detected by the weekly WP simulations. As such, the free-inquiry VP simulations appear

FIGURE 3: Comparison of Problem-Solving Index over Four Weeks of Problem-Based Learning Curriculum (PBLC) for Written and Virtual Patient Simulations (PS)



to be a more sensitive measure of CR proficiency compared to the WP simulation assessments. Interestingly, expected gains in clinical reasoning competency outcomes over the four weeks of the PBL curriculum were also detected by VP simulations but not WP simulation assessments—again supporting VP simulations as more sensitive in measuring improvements in CR competency.

It should be noted that the WP simulations were limited to measuring CR proficiency and not the other three CR constructs of efficiency, thoroughness and strategy measured by the VP simulations. It is understandable that a significant PBLC week effect might not be detected by the VP pretest if CR proficiency, efficiency and strategy skills were not being taught in the PBL curriculum. However, the first posttest (Case 5: Pulmonary) likewise did not detect a PBLC week effect for proficiency and efficiency constructs, but did detect a strong pretest-posttest VP simulation practice effect where three hours of VP practice significantly improved CR proficiency, efficiency, thoroughness and strategy scores (See Table 2). Improvements in the second posttest (Case 6: Cardiology) and not the pretest (Case 1: Cardiology) suggest a strong combined effect of the PBL curriculum and the VP elective that does not appear to be case content related as both the pretest and second posttest were from the same content domain of cardiology.

These results demonstrate the utility of free-inquiry virtual patient simulations in assessing clinical reasoning competence as the second VP simulation posttest detected expected improvements in clinical reasoning proficiency, efficiency, thoroughness, and strategy over the four-week PBL curriculum. The four clinical reasoning constructs each provide a unique view; and, the results of this study suggest that they should be considered together in assessing overall clinical reasoning competency. For example, a thorough and complete medical record is desired whenever possible, and high scores along the thoroughness construct would seem to be desirable. However, when time is limited, physicians must be able to discriminate between critical and non-critical information, and high thoroughness scores may also reflect indiscriminate data gathering and listing of medical problems as potential diagnoses.[24] Such problem solvers would tend to score low along the efficiency construct. In this example, the thoroughness and efficiency constructs taken together provide a more accurate view of overall clinical reasoning competency.

Clinical reasoning allows physicians to move from positions of clinical uncertainty to points where the medical literature can offer guidance.[25] Greater understanding of the clinical reasoning process can potentially improve patient care by helping medical students and clinicians recognize the cognitive processes underlying their decision-making errors. Both data gathering and data integration have been found to be sources of error in diagnostic decision making.[26] Computer-based diagnostic decision consultation has been found to positively influence diagnostic decision making in clinicians and students, with a larger impact upon students.[27-28] Life-sized and web-based computer simulations are also gaining increasing acceptance in PBL curricula, and may be useful in assessing clinical skills and diagnostic and therapeutic decision making.[15,29]

This study has limitations. It was conducted nearly three decades ago as part of a PhD dissertation,[30] and was never formally published. With recent developments in the ACGME core competencies,[11] new accreditation system (NAS) and milestones,[1] the a priori development and validation of CR constructs with a scoring rubric using free-inquiry VP simulations has greater relevance now than thirty years ago. VP simulations have changed in some ways that might impact study results. Still, similar or greater limitations exist for live simulated patients and the USMLE® step 3® computer-based exam.[31] Although this study was conducted at a single institution at a single point in time some years ago, more than half of a large medical school class participated, and results would be expected to generalize to other post-second year medical students with similar aptitudes and experiences. It is less clear whether results would generalize to medical students in their clinical years or to residents and physicians.

In summary, medical schools and residency programs are undergoing major transformations to ensure physician competency through outcomes measurement. This study demonstrated that even a four-week PBL curriculum can significantly impact acquisition of clinical skill and diagnostic decision-making competency, and that adding just three hours of virtual patient practice significantly augments that effect. The instructional (construct) validity of the nineteen clinical reasoning proficiency, efficiency, thoroughness, and strategy indices[20] is again suggested as they detected expected changes in clinical reasoning competencies over the four-week PBL curriculum. It seems clear that free-inquiry virtual patient simulations will have an increasingly important role in clinical reasoning outcomes assessment in the future.

Received for publication February 20, 2013. Revisions received: February 26, 2013. Accepted: February 26, 2013. Available online: February 27, 2013.

Presented in part at the Society for Academic Emergency Medicine (SAEM) 25th Annual Meeting, San Antonio, Texas, May 21-24, 1995; American Educational Research Association (AERA) Education in the Professions Annual Meeting, New York City, NY, April 10, 1996.

Contributions of Authors: Conception (DMC, JGC, APV, WKD), design (DMC, JGC), analysis and interpretation (DMC, JGC), drafting (DMC, JGC, WKD), and revising (DMC) the article. DMC takes responsibility for the paper as a whole.

Acknowledgements: The authors are indebted to Robert H. Bartlett, MD and Richard G. Judge, MD, course directors of the Problem-Based Learning Curriculum, for their enthusiastic endorsement of the medical problem solving computer simulation elective; the ninety-seven University of Michigan medical students who enrolled in the elective; the Taubman Medical Library Learning Resource Center staff for their essential role in making the computer simulation elective an overwhelming success, namely: Cindy A. McBride, Patricia Tomlin, Ellen S. Hoffman, Achla B. Karnani, Chris M. Chapman, Denise M. Axelrod, Patricia W. Martin, Dennis M. Poupart, Jackie E. McKenzie, Lois L. Katon, Mark D. Heidt, William A. Vicini, and Sharon J. Love; and, Fredric M. Wolf, PhD, Elaine M. Hockman, PhD, Nancy P. Allen, PhD, C. Victor Bunderson, PhD, Russell T. Osguthorpe, PhD and Sally H. Cavanaugh, PhD for their insightful review and recommendations to improve this work; and, to Becky Bluett and Christine Downs for their assistance in preparing the manuscript.

Funding: In part by a University of Michigan curriculum innovation grant.

Conflicts of Interest: No personal, commercial, political, academic or financial conflict of interest was reported by any of the authors. DMC is editor-in-chief of JCRPC.

Ethical Approval: Authors report the study qualified for IRB exemption.

Correspondence: Dane M. Chapman, MD, PhD, Director, Center for Clinical Reasoning and Procedural Competency, Department of Emergency Medicine, University of Missouri-Columbia School of Medicine, 1 Hospital Drive, Columbia MO, 65201, e-mail: <chapmandan@health.missouri.edu>

References

1. Nasca TJ, Philibert I, Bringham T, Flynn TC. The next GME accreditation system-Rationale and benefits. NEJM Special Report. Available online: http://www.acgme-nas.org/nejm-report.html (accessed 26 February 2013).
2. Chapman DM, Hayden S, Sanders AB et al. Integrating the Accreditation Council for Graduate Medical Education core competencies into the Model of the Clinical Practice of Emergency Medicine. *Acad Emerg Med.* 2004;11(6):674-685.
3. Van Gessel E, Nendaz MR, Vermeulen B Junod A, Vu NV. Development of clinical reasoning from the basic sciences to the clerkships: a longitudinal assessment of medical students' needs and self-perception after a transitional learning unit. *Med Educ.* 2003;37:966-974.
4. Vernon DTA, Blake RL. Does problem-based learning work? A meta-analysis of evaluative research. *Acad Med.* 1993;68:550-563.
5. Elstein AS. Clinical problem solving and decision psychology: Comment on "The epistemology of clinical reasoning". Acad Med. 2002;75-S134-6.
6. Neufeld VR, Barrows HS. The "McMaster philosophy": an approach to medical education. *J Med Educ.* 1974;49:1040-50.
7. Kinkade S. A snapshot of the status of problem-based learning in U.S. Medical Schools, 2003-2004. *Acad Med.* 2005;80:300-301.
8. Albanese MA, Mitchell S. Problem-based learning: a review of literature on its outcomes and implementation issues. *Acad Med.* 1993 Jan;68(1):52-81.
9. Hoffman K et al. Problem-Based Learning Outcomes: Ten Years of Experience at the University of Missouri-Columbia School of Medicine. *Acad Med.* 2006 July;81(7):617-25.
10. Epstein RM, Hundert EM. Defining and assessing professional competence. *JAMA* 2002;287(2):226-235.
11. Swing SR. Assessing the ACGME general competencies: general

considerations and assessment methods. *Acad Emerg Med.* 2002 Nov;9(11):1278-88.

12. ACGME/ABMS Joint Initiative: Attachment/Toolbox of Assessment Methods. Version 1.1. September 2000. Available online: http://www.partners.org/Assets/Documents/Graduate-Medical-Education/ToolTable.pdf (accessed 25 February 2013).
13. Green ML, Holmboe E. Perspective: the ACGME toolbox: half empty or half full? *Acad Med.* 2010 May;85(5):787-90.
14. Lurie SJ, Mooney CJ, Lyness JM. Commentary: pitfalls in assessment of competency-based educational objectives. *Acad Med.* 2011 Apr;86(4):412-14.
15. Murray D, Boulet J, Ziv A, Woodhouse J, Kras J, McAllister J. An acute care skills evaluation for graduating medical students: a pilot study using clinical simulation. *Med Educ.* 2002;36:833-841.
16. Friedman CP. The marvelous medical education machine or how medical education can be unstuck in time. *Acad Med.* 2000;75(10):S137-142.
17. Harless W. CASE: A computer-assisted simulation of the clinical encounter. *J of Med Educ.* 1971;46:443-448.
18. Hawkins R, Gaglione MM, LaDuca T, Leung C, Sample L, Gliva-McConvey G, Liston W, De Champlain A, Ciccone A. Assessment of patient management skills and clinical skills of practicing doctors using computer-based case simulations and standardized patients. *Med Educ.* 2004;38:958-968.
19. Campbell DT, Stanley JC 1963. E*xperimental and Quasi-Experimental Designs for Research*. Chicago: Rand McNally College Publishing Company.
20. Chapman DM, Calhoun JG, Van Mondfrans AP, Davis WK. Measuring clinical reasoning competency using a virtual patient model. *JCRPC.* 2013;1(1):1-9.
21. The University of Michigan Medical School Bulletin. 1982; pp. 11-14.
22. Wolf FM et al. Concurrent and criterion-referenced validity of patient management problems. Proceedings of the Twenty-Second Annual Conference on Research in Medical Education, 1983. Washington, D.C.: Association of American Medical Colleges, 115-121.
23. Allen NP et al. The development of patient managment problems to assess clinical problem-solving knowledge and abilities of second-year medical students (Abstract). 1982. Innovations in Medical Education Exhibits. Washington, D.C.: Association of American Medical Colleges, p. 6.
24. Voytovich AE, Rippey RM, Copertino L. Scorable problem lists as measures of clinical judgment. *Eval Health Prof.* 1980;3:159-171.
25. Kempainen RR, Migeon MB, Wolf FM. Understanding our mistakes: a primer on errors in clinical reasoning. *Med Teach.* 2003;25(2):177-181.
26. Gruppen LD, Wolf FM, Billi JE. Information gathering and integration as sources of error in diagnostic decision making. *Med Decis Making.* 1991;11(4):233-9.
27. Berner ES, Webster GD, Shugerman AA et al. Performance of four computer-based diagnostic systems. *NEJM.* 1994;330(25);1792-96.
28. Friedman CP, Elstein AS, Wolf FM et al. Enhancement of clinicians' diagnostic reasoning by computer-based consultation. *JAMA.* 1999;282:1851-56.
29. Kamin C, O'Sullivan P, Deterding R, Younger M. A comparison of critical thinking in groups of third-year medical students in text, video, and virtual PBL case modalities. *Acad Med.* 2003;78(2):204-211.
30. Chapman DM. Teaching and evaluating clinical reasoning through computer-based patient management simulations. Dissertation Abstracts International. 1985;46:784-B.
31. USMLE Bulletin of Information. Available online: http://www.usmle.org/bulletin/overview (accessed 25 February 2013).

Emergency Cricothyrotomy

Video Clip Demonstration

DANE M. CHAPMAN, MD, PHD,[1,2] JOHN A. MARX, MD,[1,3]
BENJAMIN HONIGMAN, MD,[1] PETER ROSEN, MD[1,4]

Institutional Affiliations:

1. *University of Colorado-Denver School of Medicine and Denver Health Medical Center, Departments of Emergency Medicine, Denver, CO*
2. *University of Missouri-Columbia School of Medicine, Department of Emergency Medicine, Columbia, MO*
3. *Carolinas Medical Center, Department of Emergency Medicine, Charlotte, NC (Deceased)*
4. *Harvard Medical School and Beth Israel Deaconess Medical Center, Department of Emergency Medicine, Boston, MA.*

During the development of one of the first emergency procedure-based computer simulations,[1-4] videotaping procedure demonstrations by experts was required. The mastered videotape was then reproduced using what was then considered state-of-the-art videodisc technology (precursor to the CD-ROM and DVD). Complex video images on the videodisc were then programmed into an interactive videodisc-computer simulation designed to teach and assess critical emergency procedural skills.[1-4] This development effort was funded in part by several organizations[a] and resulted in many hours of interactive procedural skill instruction using the videodisc-computer platform. As the videodisc was available only to participating medical schools of the Health Science Consortium,[4] it was not widely distributed. In an effort to more broadly disseminate the instructional portion of the videodisc, emergency procedures from the mastered videotape were made available for wider distribution using the CD-ROM format.[5] Now, for the first time, video clips of these critical emergency procedural skill demonstrations are being made available online.

[a]*Funding:* Emergency Medicine Foundation, Resident Research Award, 7/88 - 12/89; Teaching Enrichment Grant, University of Colorado School of Medicine, 1/89 - 1/90; Subcomponent of Institutional IBM Infowindow Consortium Grant, University of Colorado Health Sciences Center, 1/89 - 1/90; Society of Academic Emergency Medicine and Emergency Medicine Foundation, Educational Methodology Grant, 11/90 - 1/94.

Previously Published in Part: Emergency Medicine on CD-ROM, Lippincott-Raven Publishers, 1997.

References

1. Chapman DM, Rhee KJ, Marx JA, Honigman B, Panacek EA, Martinez D, Brofeldt BT, Cavanaugh SH. Open thoracotomy procedural competency: A validity study of teaching and assessment modalities. *Ann Emerg Med.* December 1996;28:641-647.
2. Chapman DM, Marx JA, Honigman B, Rosen P, Cavanagh, SH. Emergency thoracotomy: A comparison of medical student, resident, and faculty performances on written, computer, and animal model assessments. *Acad Emerg Med.* 1994;1:373-381.
3. Chapman DM. Use of computer-based technologies in teaching emergency procedural skills. *Acad Emerg Med.* 1994;1:404-407.
4. O'Neil PN. Developing video instruction for the health sciences: A consortial approach. *Acad Med.* 1990;65:624-627.
5. Chapman DM, Marx JA, Honigman B, and Rosen P. Emergency cricothyrotomy (video clip demonstration) in *Emergency Medicine on CD-ROM*, Lippincott-Raven Publishers, 1997.

Emergency Thoracotomy

Video Clip Demonstration

DANE M. CHAPMAN MD, PHD,[1,2] JOHN A. MARX MD,[1,3]
BENJAMIN HONIGMAN, MD,[1] PETER ROSEN, MD[1,4]

Institutional Affiliations:

1. *University of Colorado-Denver School of Medicine and Denver Health Medical Center, Departments of Emergency Medicine, Denver, CO*
2. *University of Missouri-Columbia School of Medicine, Department of Emergency Medicine, Columbia, MO*
3. *Carolinas Medical Center, Department of Emergency Medicine, Charlotte, NC (Deceased)*
4. *Harvard Medical School and Beth Israel Deaconess Medical Center, Department of Emergency Medicine, Boston, MA.*

During the development of one of the first emergency procedure-based computer simulations,[1-4] videotaping procedure demonstrations by experts was required. The mastered videotape was then reproduced using what was then considered state-of-the-art videodisc technology (precursor to the CD-ROM and DVD). Complex video images on the videodisc were then programmed into an interactive videodisc-computer simulation designed to teach and assess critical emergency procedural skills.[1-4] This development effort was funded in part by several organizations[a] and resulted in many hours of interactive procedural skill instruction using the videodisc-computer platform. As the videodisc was available only to participating medical schools of the Health Science Consortium,[4] it was not widely distributed. In an effort to more broadly disseminate the instructional portion of the videodisc, emergency procedures from the mastered videotape were made available for wider distribution using the CD-ROM format.[5] Now, for the first time, video clips of these critical emergency procedural skill demonstrations are being made available online.

---

[a]*Funding:* Emergency Medicine Foundation, Resident Research Award, 7/88 - 12/89; Teaching Enrichment Grant, University of Colorado School of Medicine, 1/89 - 1/90; Subcomponent of Institutional IBM Infowindow Consortium Grant, University of Colorado Health Sciences Center, 1/89 - 1/90; Society of Academic Emergency Medicine and Emergency Medicine Foundation, Educational Methodology Grant, 11/90 - 1/94.

Previously Published in Part: Emergency Medicine on CD-ROM, Lippincott-Raven Publishers, 1997.

References

1. Chapman DM, Rhee KJ, Marx JA, Honigman B, Panacek EA, Martinez D, Brofeldt BT, Cavanaugh SH. Open thoracotomy procedural competency: A validity study of teaching and assessment modalities. *Ann Emerg Med.* December 1996;28:641-647.
2. Chapman DM, Marx JA, Honigman B, Rosen P, Cavanagh, SH. Emergency thoracotomy: A comparison of medical student, resident, and faculty performances on written, computer, and animal model assessments. *Acad Emerg Med.* 1994;1:373-381.
3. Chapman DM. Use of computer-based technologies in teaching emergency procedural skills. *Acad Emerg Med.* 1994;1:404-407.
4. O'Neil PN. Developing video instruction for the health sciences: A consortial approach. *Acad Med.* 1990;65:624-627.
5. Chapman DM, Marx JA, Honigman B, and Rosen P. Emergency thoracotomy (video clip demonstration) in *Emergency Medicine on CD-ROM*, Lippincott-Raven Publishers, 1997.

Learning Preference Inventory

As part of a PhD dissertation in medical education, permission was obtained from Agnes G. Rezler, PhD to administer the Learning Preference Inventory (LPI)[1] to medical students at the University of Michigan Medical School.[2] That work demonstrated the usefulness of the LPI as a valid and reliable measure of medical students who would most benefit from problem-based and virtual patient simulation curricula; and was formally published in 2006.[3]

Since 2006, the authors have been contacted by other investigators in their efforts to obtain copies of the Learning Preference Inventory. To our knowledge, the LPI is no longer available.

In honor of Agnes G. Rezler, PhD (b. August 26, 1922 d. April 8, 2001), we publish the Learning Preference Inventory in this Inaugural Issue of the Journal. Dr. Rezler spent much of her professional career developing and validating the LPI and we feel she would be pleased to know that her work continues on in a useful manner for the benefit of medical education. It is reproduced here in its original format (Appendix A) and digitalized as a survey ready for distribution (Appendix B). The scoring sheet and methodology are also included.

References

1. Rezler AG, Rezmovic V. The Learning Preference Inventory. *J Allied Health.* 1981:10:28-34.
2. Personal correspondence from Agnes G. Rezler, PhD dated April 11, 1984.
3. Chapman DM, Calhoun JG. Validation of learning style measures: Implications for medical education practice. *Med Educ.* 2006; 40:576-583.

LEARNING PREFERENCE INVENTORY (LPI)

This Inventory gives you the chance to indicate your method of learning. It is not a "test"; there are no right or wrong answers. The aim of the Inventory is to describe how you learn, not to evaluate your learning ability.

The Inventory has two parts. In Part I there are six sets of six words listed. In Part II there are nine items, each of which contain six statements.

Instructions for answering Part I

1. Record all of your answers on the Answer Sheet for Part I; make no marks on the Inventory itself.
2. Print your last name and initials on the space provided on the Answer Sheet for Part I. Circle the number corresponding to your profession: circle 1 or 2 for faculty or student, and 1 or 2 for male or female.
3. Read all six colors shown in the left hand column:

a.	yellow	a. 6
b.	green	b. 3
c.	blue	c. 4
d.	red	d. 5
e.	white	e. 1
f.	black	f. 2

The above listed colors are ranked in order of preference in the right hand column. This ranking shows that yellow is the most preferred color (6) and white is the least preferred color (1).

Now read all words listed in Part I, Columns A through F. Rank all six words in each column according to your learning perferences: Write 6 on the Answer Sheet for the word that promotes learning most for you and 1 next to the word that promotes learning least for you. Assign numbers 2, 3, 4 and 5 to the remaining words in each column.

Rank EACH word: please do not omit words. Be sure to assign a different rank to each of the six words in each column; do not make ties.

Request copies from: Agnes G. Rezler, Ph.D., University of Illinois, at the Medical Center, C.E.D., 808 South Wood Street, Chicago, Illinois 60612

LEARNING PREFERENCE INVENTORY

Part I

6 = promotes learning <u>most for you</u>
5 = promotes learning <u>second best</u>
4 = promotes learning third best
3 = promotes learning fourth best
2 = promotes learning fifth best
1 = promotes learning <u>least for you</u>

COLUMN A	COLUMN B	COLUMN C
a. factual	a. self-instructional	a. sharing
b. teacher directed	b. myself	b. doing
c. teamwork	c. hypothetical	c. guided
d. reading	d. interpersonal	d. self-initiated
e. self-evaluation	e. teacher-defined	e. thinking
f. theoretical	f. practical	f. solitary

COLUMN D	COLUMN E	COLUMN F
a. teacher-structured	a. scientific	a. individual
b. concrete	b. assigned	b. applied
c. writing	c. skill-oriented	c. supervised
d. group	d. personal	d. autonomous
e. conceptual	e. self-designed	e. abstract
f. self-directed	f. team-oriented	f. interactive

Instructions for answering Part II

1. Record all of your answers on the Answer Sheet for Part II; make no marks on the Inventory itself.

2. Read the first item in Part II and rank order all six responses. Write 6 for the statement in Item I which promotes learning most for you, and 1 for the statement that promotes learning least for you. Assign numbers 2, 3, 4 and 5 to the remaining statements in each item.

   Be sure to assign a different ~~rank~~ number to each of the six statements in Item I and continue the same procedure with all nine items. Rank each statement; please do not omit statements and do not make ties.

LEARNING PREFERENCE INVENTORY

Part II

6 = promotes learning <u>most for you</u>
5 = promotes learning second best
4 = promotes learning third best
3 = promotes learning fourth best
2 = promotes learning fifth best
1 = promotes learning <u>least for you</u>

I. Read the following six statements and then rank them in terms of how well they describe the teachers in whose classes you enjoyed learning.

a. The teacher gave many practical, concrete examples.
b. The teacher let me set my own goals.
c. The teacher encouraged me to work by myself.
d. The teacher was friendly and outgoing.
e. The teacher made the relationships between different schools of thought clear.
f. The teacher made clear and definite assignments and I knew exactly what was expected.

II. Number the following kinds of work in the order in which they would interest you.

a. Work that would require cooperation among team members.
b. Work with specific and practical ways of handling things.
c. Work that would let me do things on my own.
d. Work that would permit me to deal with ideas rather than things.
e. Work that I could plan and organize myself.
f. Work that would be clearly defined and specified by my supervisor.

III. Rank the following in terms of their effect on how hard you work and how much you accomplish in a class.

a. I can set my own goals and proceed accordingly.
b. I can address myself to a concrete, practical task.
c. I have an opportunity to discuss or work on something with other students.
d. I can examine different schools of thought.
e. I understand what is expected, when work is due and how it will be evaluated.
f. I can accomplish most tasks by myself.

IV. The evaluation of student performance is a part of nearly all courses. Rank the following in terms of how you feel about such evaluation.

a. It should be assembled from questions provided by students.
b. It should focus on individual performance.
c. It should consist of a written examination dealing mainly with basic concepts.
d. It should consist of a practical examination dealing with skills.
e. It should be consistent with clearly specified requirements.
f. It should not interfere with good relationships between the teacher and the student.

6 = promotes learning most for you
5 = promotes learning second best
4 = promotes learning third best
3 = promotes learning fourth best
2 = promotes learning fifth best
1 = promotes learning least for you

V. Rank the following in terms of their general value to you as you learn.

a. Study a textbook.
b. Engage in an internship or practicum.
c. Prepare a class project with other students.
d. Search for reasons to explain occurrences.
e. Follow a prepared outline by the teacher.
f. Prepare your own outline.

VI. Rank the following in terms of how much they would attract you to an elective class.

a. Good personal relationships between teacher and students
b. Clearly spelled out standards and requirements.
c. Emphasis on practicing skills
d. Emphasis on independent study.
e. Opportunity to determine own activities.
f. Emphasis on theoretical concepts.

VII. Consider the following in terms of their general effect on how well you do in a class.

a. I can study on my own.
b. I can work with something tangible
c. I can focus on ideas and concepts.
d. I can organize things my own way.
e. I can work with others.
f. I can work on clear-cut assignments.

VIII. Rank the following in the order in which you think teachers should possess these characteristics or skills.

a. Getting students to set their own goals.
b. Getting students to demonstrate concrete skills.
c. Involving students in generating hypotheses.
d. Preparing self-instructional materials.
e. Relating well to students.
f. Planning all aspects of courses and learning activities.

IX. Rank the following in terms of how much they generally help you learn and remember.

a. Studying alone instead of studying with fellow students.
b. Performing a specific task.
c. Having a knowledgeable teacher discuss theory upon which practice is built.
d. Determining your own approach and proceeding accordingly.
e. Joining a student group to study together and share ideas.
f. Getting an outline of the course from the teacher and a clear understanding of what will occur in the course.

LEARNING PREFERENCE INVENTORY

Answer Sheet

Part I

(1-13) Last Name and Initials ____________________

(14) Profession:(circle one or none)	1. dentistry	4. pharmacy
	2. nursing	5. public health
	3. medicine	6. behavior science
(15) Profession:(circle one or none)	1. dietetics	4. medical technology
	2. occupational therapy	5. medical records
	3. physical therapy	6. x-ray technology
		7. other, specify ____________

(16) Circle one: 1. faculty 2. student (17) Circle one: 1. male 2. female

(18) Card Number 1

COLUMN A	COLUMN B	COLUMN C
(26) a._____	(32) a._____	(38) a._____
(27) b._____	(33) b._____	(39) b._____
(28) c._____	(34) c._____	(40) c._____
(29) d._____	(35) d._____	(41) d._____
(30) e._____	(36) e._____	(42) e._____
(31) f._____	(37) f._____	(43) f._____

COLUMN D	COLUMN E	COLUMN F
(44) a._____	(50) a._____	(56) a._____
(45) b._____	(51) b._____	(57) b._____
(46) c._____	(52) c._____	(58) c._____
(47) d._____	(53) d._____	(59) d._____
(48) e._____	(54) e._____	(60) e._____
(49) f._____	(55) f._____	(61) f._____

(OVER)

LEARNING PREFERENCE INVENTORY

Answer Sheet

Part II

(18) Card Number __2__

ITEM I	ITEM II	ITEM III	ITEM IV	ITEM V
(26) a.____	(32) a.____	(38) a.____	(44) a.____	(50) a.____
(27) b.____	(33) b.____	(39) b.____	(45) b.____	(51) b.____
(28) c.____	(34) c.____	(40) c.____	(46) c.____	(52) c.____
(29) d.____	(35) d.____	(41) d.____	(47) d.____	(53) d.____
(30) e.____	(36) e.____	(42) e.____	(48) e.____	(54) e.____
(31) f.____	(37) f.____	(43) f.____	(49) f.____	(55) f.____

ITEM VI	ITEM VII	ITEM VIII	ITEM IX
(56) a.____	(62) a.____	(68) a.____	(74) a.____
(57) b.____	(63) b.____	(69) b.____	(75) b.____
(58) c.____	(64) c.____	(70) c.____	(76) c.____
(59) d.____	(65) d.____	(71) d.____	(77) d.____
(60) e.____	(66) e.____	(72) e.____	(78) e.____
(61) f.____	(67) f.____	(73) f.____	(79) f.____

Scoring Your Learning Preference Inventory

You may utilize this sheet to calculate your scores on the different subscales of the LPI. Each number in Part I and II listed below corresponds to an item on the LPI. For each of the items, write the rank that you gave it on the LPI answer sheet (remember, 6 represents the highest rank and 1, the lowest). After writing in the appropriate ranks, total them separately for Parts I and II. At the bottom of the page a space is provided for you to combine the totals of both parts.

Part I

AB	CO	TS	SS	IP	IN
(31)____	(26)____	(27)____	(30)____	(28)____	(29)____
(34)____	(37)____	(36)____	(32)____	(35)____	(33)____
(42)____	(39)____	(40)____	(41)____	(38)____	(43)____
(48)____	(45)____	(44)____	(49)____	(47)____	(46)____
(50)____	(52)____	(51)____	(54)____	(55)____	(53)____
(60)____	(57)____	(58)____	(59)____	(61)____	(56)____

art I
ubtotal: ____ + ____ + ____ + ____ + ____ + ____ = 126

Part II

(30)____	(26)____	(31)____	(27)____	(29)____	(28)____
(35)____	(33)____	(37)____	(36)____	(32)____	(34)____
(41)____	(39)____	(42)____	(38)____	(40)____	(43)____
(46)____	(47)____	(48)____	(44)____	(49)____	(45)____
(53)____	(51)____	(54)____	(55)____	(52)____	(50)____
(61)____	(58)____	(57)____	(60)____	(56)____	(59)____
(64)____	(63)____	(67)____	(65)____	(66)____	(62)____
(70)____	(69)____	(73)____	(68)____	(72)____	(71)____
(76)____	(75)____	(79)____	(77)____	(78)____	(74)____

art II
ubtotal: ____ + ____ + ____ + ____ + ____ + ____ = 189

otals ____ + ____ + ____ + ____ + ____ + ____ = 315
Add both subtotals in each column)

Adapted From: Foley, R.P. and Smilansky, J. Teaching Techniques: A Handbook for Health Professionals. McGraw-Hill: New York, 1980, p. 108.

LEARNING PREFERENCE INVENTORY (LPI)

This Inventory gives you the chance to indicate your method of learning. It is not a "test"; there are no right or wrong answers. The aim of the Inventory is to describe how you learn, not to evaluate your learning ability. The Inventory has two parts. In Part I there are six sets of six words listed. In Part II there are nine items, each of which contain six statements.

PART I

Instructions for Answering Part I

1. Record all of your answers on the Answer Sheet for Part I; make no marks on the Inventory itself.
2. Print your last name and initials on the space provided on the Answer Sheet for Part I. Circle the number corresponding to your profession: circle 1 or 2 for faculty or student, and 1 or 2 for male or female.
3. As an illustration of ranking procedure, read all six colors shown in the left hand column:

a.	yellow	a.	6
b.	green	b.	3
c.	blue	c.	4
d.	red	d.	5
e.	white	e.	1
f.	black	f.	2

The above listed colors are ranked in order of preference in the right hand column. This ranking shows that yellow is the most preferred color (6) and white is the least preferred color (1).

Now, read all the words listed below in Part I, Columns A through F. Rank all six words in each column according to your learning preferences: write 6 on the Answer Sheet for the word that promotes learning most for you and 1 next to the word that promotes learning least for you. Assign numbers 2, 3, 4 and 5 to the remaining words in each column.

Rank EACH word; please do not omit words. Be sure to assign a different rank to each of the six words in each column; do not make ties.

Adapted From: Foley RP, Smilansky J. *Teaching Techniques: A Handbook for Health Professionals*. McGraw-Hill: New York, 1980:108.

6 = promotes learning most for you
5 = promotes learning second best
4 = promotes learning third best
3 = promotes learning fourth best
2 = promotes learning fifth best
1 = promotes learning least for you

COLUMN A

a. factual
b. teacher directed
c. teamwork
d. reading
e. self-evaluation
f. theoretical

COLUMN B

a. self-instructional
b. myself
c. hypothetical
d. interpersonal
e. teacher-defined
f. practical

COLUMN C

a. sharing
b. doing
c. guided
d. self-initiated
e. thinking
f. solitary

COLUMN D

a. teacher-structured
b. concrete
c. writing
d. group
e. conceptual
f. self-directed

COLUMN E

a. scientific
b. assigned
c. skill-oriented
d. personal
e. self-designed
f. team-oriented

COLUMN F

a. individual
b. applied
c. supervised
d. autonomous
e. abstract
f. interactive

PART II

Instructions for Answering Part II

1. Record all of your answers on the Answer Sheet for Part II; make no marks on the Inventory itself.

2. Read the first item in Part II and rank order all six responses. Write 6 for the statement in Item I which promotes learning most for you, and 1 for the statement that promotes learning least for you. Assign numbers 2, 3, 4 and 5 to the remaining statements in each item.

   Be sure to assign a different rank (number) to each of the six statements in Item I and continue the same procedure with all nine items. Rank each statement; please do not omit statements and do not make ties.

6 = promotes learning most for you
5 = promotes learning second best
4 = promotes learning third best
3 = promotes learning fourth best
2 = promotes learning fifth best
1 = promotes learning least for you

I. Read the following six statements and then rank them in terms of how well they describe the teachers in whose classes you enjoyed learning.

a. The teacher gave many practical, concrete examples.
b. The teacher let me set my own goals.
c. The teacher encouraged me to work by myself.
d. The teacher was friendly and outgoing.
e. The teacher made the relationships between different schools of thought clear.
f. The teacher made clear and definite assignments and I knew exactly what was expected.

II. Number the following kinds of work in the order in which they would interest you.

a. Work that would require cooperation among team members.
b. Work with specific and practical ways of handling things.
c. Work that would let me do things on my own.
d. Work that would permit me to deal with ideas rather than things.
e. Work that I could plan and organize myself.
f. Work that would be clearly defined and specified by my supervisor.

III. Rank the following in terms of their effect on how hard you work and how much you accomplish in a class.

a. I can set my own goals and proceed accordingly.
b. I can address myself to a concrete, practical task.
c. I have an opportunity to discuss or work on something with other students.
d. I can examine different schools of thought.
e. I understand what is expected, when work is due and how it will be evaluated.
f. I can accomplish most tasks by myself.

IV. The evaluation of student performance is a part of nearly all courses. Rank the following in terms of how you feel about such evaluation.

a. It should be assembled from questions provided by students.
b. It should focus on individual performance.
c. It should consist of a written examination dealing mainly with basic concepts.
d. It should consist of a practical examination dealing with skills.
e. It should be consistent with clearly specified requirements.
f. It should not interfere with good relationships between the teacher and the student.

6 = promotes learning most for you
5 = promotes learning second best
4 = promotes learning third best
3 = promotes learning fourth best
2 = promotes learning fifth best
1 = promotes learning least for you

V. Rank the following in terms of their general value to you as you learn.

a. Study a textbook.
b. Engage in an internship or practicum.
c. Prepare a class project with other students.
d. Search for reasons to explain occurrences.
e. Follow a prepared outline by the teacher.
f. Prepare your own outline.

VI. Rank the following in terms of how much they would attract you to an elective class.

a. Good personal relationships between teacher and students
b. Clearly spelled out standards and requirements.
c. Emphasis on practicing skills
d. Emphasis on independent study.
e. Opportunity to determine own activities.
f. Emphasis on theoretical concepts.

VII. Consider the following in terms of their general effect on how well you do in a class.

a. I can study on my own.
b. I can work with something tangible
c. I can focus on ideas and concepts.
d. I can organize things my own way.
e. I can work with others.
f. I can work on clear-cut assignments.

VIII. Rank the following in the order in which you think teachers should possess these characteristics or skills.

a. Getting students to set their own goals.
b. Getting students to demonstrate concrete skills.
c. Involving students in generating hypotheses.
d. Preparing self-instructional materials.
e. Relating well to students.
f. Planning all aspects of courses and learning activities.

IX. Rank the following in terms of how much they generally help you learn and remember.

a. Studying alone instead of studying with fellow students.
b. Performing a specific task.
c. Having a knowledgeable teacher discuss theory upon which practice is built.
d. Determining your own approach and proceeding accordingly.
e. Joining a student group to study together and share ideas.
f. Getting an outline of the course from the teacher and a clear understanding of what will occur in the course.

LEARNING PREFERENCE INVENTORY

Answer Sheet – Part I

(1-13) Last Name and Initials ______________________

(14) Profession: (circle one or none)

1. Dentistry
2. Nursing
3. Medicine
4. Pharmacy
5. Public Health
6. Behavior Science

(15) Profession: (circle one or none)

1. Dietetics
2. Occupational Therapy
3. Physical Therapy
4. Medical Technology
5. Medical Records
6. X-ray Technology
7. Other, specify ______________________

(16) Circle one: 1. Faculty 2. Student (17) Circle one: 1. Male 2. Female

(18) Card Number ____1____

COLUMN A			COLUMN B			COLUMN C		
(26)	a.	______	(32)	a.	______	(38)	a.	______
(27)	b.	______	(33)	b.	______	(39)	b.	______
(28)	c.	______	(34)	c.	______	(40)	c.	______
(29)	d.	______	(35)	d.	______	(41)	d.	______
(30)	e.	______	(36)	e.	______	(42)	e.	______
(31)	f.	______	(37)	f.	______	(43)	f.	______

COLUMN D			COLUMN E			COLUMN F		
(44)	a.	______	(50)	a.	______	(56)	a.	______
(45)	b.	______	(51)	b.	______	(57)	b.	______
(46)	c.	______	(52)	c.	______	(58)	c.	______
(47)	d.	______	(53)	d.	______	(59)	d.	______
(48)	e.	______	(54)	e.	______	(60)	e.	______
(49)	f.	______	(55)	f.	______	(61)	f.	______

LEARNING PREFERENCE INVENTORY
Answer Sheet – Part II

(18) Card Number ____2____

ITEM I	ITEM II	ITEM III	ITEM IV	ITEM V
(26) a. ______	(32) a. ______	(38) a. ______	(44) a. ______	(50) a. ______
(27) b. ______	(33) b. ______	(39) b. ______	(45) b. ______	(51) b. ______
(28) c. ______	(34) c. ______	(40) c. ______	(46) c. ______	(52) c. ______
(29) d. ______	(35) d. ______	(41) d. ______	(47) d. ______	(53) d. ______
(30) e. ______	(36) e. ______	(42) e. ______	(48) e. ______	54) e. ______
(31) f. ______	(37) f. ______	(43) f. ______	(49) f. ______	(55) f. ______

ITEM VI	ITEM VII	ITEM VIII	ITEM IX
(56) a. ______	(62) a. ______	(68) a. ______	(74) a. ______
(57) b. ______	(63) b. ______	(69) b. ______	(75) b. ______
(58) c. ______	(64) c. ______	(70) c. ______	(76) c. ______
(59) d. ______	(65) d. ______	(71) d. ______	(77) d. ______
(60) e. ______	(66) e. ______	(72) e. ______	(78) e. ______
(61) f. ______	(67) f. ______	(73) f. ______	(79) f. ______

Scoring Your Learning Preference Inventory

You may utilize this sheet to calculate your scores on the different subscales of the LPI. Each number in Part I and II listed below corresponds to an item on the LPI. For each of the items, write the rank that you gave it on the LPI answer sheet (remember, 6 represents the highest rank and 1, the lowest). After writing in the appropriate ranks, total them separately for Parts I and II. At the bottom of the page a space is provided for you to combine the totals of both parts.

Part I

	AB		CO		TS		SS		IP		IN	
	(31)	______	(26)	______	(27)	______	(30)	______	(28)	______	(29)	______
	(34)	______	(37)	______	(36)	______	(32)	______	(35)	______	(33)	______
	(42)	______	(39)	______	(40)	______	(41)	______	(38)	______	(43)	______
	(48)	______	(45)	______	(44)	______	(49)	______	(47)	______	(46)	______
	(50)	______	(52)	______	(51)	______	(54)	______	(55)	______	(53)	______
	(60)	______	(57)	______	(58)	______	(59)	______	(61)	______	(56)	______
Part I Subtotal:	______ +		______ +		______ +		______ +		______ +		______	= 126

Part II

	(30)	______	(26)	______	(31)	______	(27)	______	(29)	______	(28)	______
	(35)	______	(33)	______	(37)	______	(36)	______	(32)	______	(34)	______
	(41)	______	(39)	______	(42)	______	(38)	______	(40)	______	(43)	______
	(46)	______	(47)	______	(48)	______	(44)	______	(49)	______	(45)	______
	(53)	______	(51)	______	(54)	______	(55)	______	(52)	______	(50)	______
	(61)	______	(58)	______	(57)	______	(60)	______	(56)	______	(59)	______
	(64)	______	(63)	______	(67)	______	(65)	______	(66)	______	(62)	______
	(70)	______	(69)	______	(73)	______	(68)	______	(72)	______	(71)	______
	(76)	______	(75)	______	(79)	______	(77)	______	(78)	______	(74)	______
Part II Subtotal:	______ +		______ +		______ +		______ +		______ +		______	= 189

Totals: ______ + ______ + ______ + ______ + ______ + ______ = 315

(Add both subtotals in each column)

Adapted From: Foley RP, Smilansky J. *Teaching Techniques: A Handbook for Health Professionals*. McGraw-Hill: New York, 1980:108.