

1996

Should We Test for Interpersonal Lawyering Skills?

Lawrence M. Grosberg

New York Law School, lawrence.grosberg@nyls.edu

Follow this and additional works at: http://digitalcommons.nyls.edu/fac_articles_chapters

Recommended Citation

2 *Clinical L. Rev.* 349 1995-1996

This Article is brought to you for free and open access by the Faculty Scholarship at DigitalCommons@NYLS. It has been accepted for inclusion in Articles & Chapters by an authorized administrator of DigitalCommons@NYLS.

SHOULD WE TEST FOR INTERPERSONAL LAWYERING SKILLS?

LAWRENCE M. GROSBERG*

INTRODUCTION

As clinicians we are engaged in preparing students to practice law. Part of that task, though certainly not all, is teaching particular lawyering skills, such as interviewing, counseling, negotiating and trial advocacy. Our work rests on a shared conviction that these skills can be performed well or badly, and that it is possible to identify at least some of the elements of superior performance and convey those to our students. Just as firmly we believe that it is possible to provide critiques of our students' efforts to acquire and use these skills.¹ Believing that skills can be identified, explained and critically analyzed, we must also believe that students' mastery of them can be evaluated—whether informally, through the feedback we provide in supervision,² or formally, through grades. Indeed, though grading is not universal in in-house clinics, it appears now to be quite common.

While we are very much immersed in the task of evaluating lawyering skills, I do not believe that we have studied as carefully as we should the question of whether there are alternative ways to accomplish those responsibilities. If anything, those of us who teach in live-client clinics³ may have come to take it for granted that the only

* Professor of Law, New York Law School. I want to thank Carol Buckler, Rick Marcico and Cynthia Schuman for their helpful comments and suggestions on earlier drafts of this article, and New York Law School for its support of my efforts to experiment with many of the ideas discussed here.

¹ See Kenneth R. Kreiling, *Clinical Education and Lawyer Competency: The Process of Learning to Learn from Experience Through Properly Structured Clinical Supervision*, 40 MD. L. REV. 284 (1981). It also is worth noting that members of the bar increasingly agree that these skills can be taught in law school. See Bryant G. Garth & Joanne Martin, *Law Schools and the Construction of Competence*, 43 J. LEGAL EDUC. 469, 479-81 (1993); Joanne Martin & Bryant G. Garth, *Clinical Education as a Bridge Between Law School and Practice: Mitigating the Misery*, 1 CLIN. L. REV. 443, 447-48 (1994).

² See, e.g., Richard K. Neumann, Jr., *A Preliminary Inquiry into the Art of Critique*, 40 HASTINGS L. J. 725 (1989); Amy L. Zeigler, *Developing a System of Evaluating in Clinical Legal Teaching*, 42 J. LEGAL EDUC. 575 (1992).

³ "Live-client" or "in-house" clinics are courses in which the students represent real clients under the direct supervision of clinical teachers who are generally full-time members of the law school faculty. The scope of students' duties and responsibilities will vary depending on the nature of the matters and the rules of the jurisdiction regarding student practice. Generally, however, the goal is to give the students as much responsibility as possible. See generally *Report of the Committee on the Future of the In-House Clinic*, 42 J. LEGAL EDUC. 508 (1992).

proper way to evaluate our students' acquisition of skills is through our ongoing, direct observation of the students' actual performances in role, either with real clients or in simulations. This assumption has some justification, for as I assert below this accumulation of one-on-one observations does seem to be the best basis for the evaluation of lawyering skills in the setting of the clinic. But it is not a perfect method, even in the clinics, and outside the clinics—in the many other settings where clinical teaching broadly understood is now underway⁴—one-on-one observation may be completely out of the question.

In these other settings, therefore, we must ask whether there is any other way to evaluate the lawyering skills being taught. The central concern of this Article is to answer this question. My answer is that students' lawyering skills can be evaluated, at least to some extent, by using any of several different kinds of tests or exams.⁵ We need to address this issue not only because we ought to be evaluating what our students have learned, but also because a failure to evaluate may ultimately compromise our success in teaching. In focusing on exams, this Article focuses on the evaluation method most relied upon in classroom legal education. Because of the impact of exams on grades, what is tested in the exams inevitably affects what is learned in the classroom. Each necessarily affects the success of the other. If skills teaching is occurring increasingly in the non-clinical classroom and *not* being tested on the final exam, the impact of this teaching may be diminished.⁶ Most disturbingly, we may even send a silent

⁴ Roleplaying, simulation and videotaped lawyering vignettes—some of the tools of clinical teaching—now find their way into many different kinds of courses, ranging from first year "Introduction to Lawyering" courses to upper class simulation courses such as "Pre-trial Advocacy" and substantive law seminars and workshops. See generally *Symposium on Simulations*, 45 J. LEGAL EDUC. 469 (1995) (including articles about the use of simulations in civil procedure, business law, corporations and pre-trial practice classes).

⁵ My interest in skills testing parallels an increased attention by law professors generally to the topic of testing and to the improvement of our methods of testing competence as a lawyer. See, e.g., Philip C. Kissam, *Law School Examinations*, 42 VAND. L. REV. 433 (1989); Norman Redlich & Steve Friedland, *Challenging Tradition: Using Objective Questions in Law School Examinations*, 41 DEPAUL L. REV. 143 (1991); Paul T. Wangerin, *Grade Conferences from Hell: Measurement Error in Law School Grading* (unpublished paper presented to Gonzaga University School of Law, Institute for Law School Teaching, Conference on "The Science and Art of Law Teaching" (July 1994)) (on file with the author). For further discussion of testing issues, see generally LUCY C. JACOBS, *DEVELOPING AND USING TESTS EFFECTIVELY: A GUIDE FOR FACULTY* (1992). The recently established Institute for Law Teaching at the Gonzaga University School of Law has now conducted workshops on testing at each of its most recent annual conferences, in 1994 and 1995. The AALS also dealt with the topic of testing both at its annual convention in 1992 and at its teaching conference in Minneapolis in June, 1995.

⁶ Myron Moskovitz, *Beyond the Case Method: It's Time to Teach with Problems*, 42 J. LEGAL EDUC. 241, 261 (1992) ("teaching and testing work best when they complement

message to our students that these untested skills are not truly important—the wrong message to be communicating to future lawyers.⁷

At the outset, it is important to define what a “test” is. Generally, “a test is a set of questions or situations designed to permit an inference about what an examinee knows or can do in an area of interest.”⁸ The test taker may be expected to state an answer in writing or orally,⁹ or may be required to “perform an act or produce a product.”¹⁰ Thus law school tests clearly can take many forms, from various types of “paper and pencil” exams¹¹ to examinations that involve the grading of actual lawyering work. So, for example, an examination in a clinic might consist of a teacher’s observing and grading the student’s trial of a real or simulated case, or the student’s conduct of a real or simulated interview. Similarly, a student might be graded on a brief or a negotiation outline he or she had prepared in an actual case. Ultimately, any occasion for evaluation might be described as a test. In the law school clinic, the clinician’s regular observation and analysis of the student’s performance is a form of evaluation or test of the student’s ability to perform any of various lawyering tasks; the cumulative result of such consecutive evaluations typically will result in a clinic course grade. We usually think of tests, however, as calling for student answers, or student performance, under special time limits

each other”); cf. Mark Tushnet, *Evaluating Students as Preparation for the Practice of Law*, 8 GEO. J. LEGAL ETHICS 313 (1995) (discussing an effort to design an evaluation system that would encourage students to work collaboratively in the classroom and in preparation for the final exam).

⁷ I agree with Professor Roy Stuckey that educational goals are not necessarily the same as testing goals, for there are some lawyering qualities, such as integrity, that we certainly value and want our students to value but probably cannot test for. But there are many lawyering attributes that can be tested, and should be. See Roy T. Stuckey, *Apprenticeships and Clinical Education: The Only Real Performance Tesis?*, BAR EXAMINER, Aug. 1986, at 7.

⁸ NATIONAL COMMISSION ON TESTING AND POLICY, FROM GATEKEEPER TO GATEWAY: TRANSFORMING TESTING IN AMERICA 2 (1990) [hereinafter FROM GATEKEEPER TO GATEWAY].

⁹ Oral exams are a customary part of the process of obtaining a doctorate in this country. Many European law schools give oral examinations in substantive law subjects. Interview with Professor Leszek Leszczynski, University of Marie Curie Skłodowska Law Faculty, in Lublin, Poland (Oct. 10, 1995).

¹⁰ FROM GATEKEEPER TO GATEWAY, *supra* note 8, at 2.

¹¹ I use the term “paper and pencil” to refer loosely to any examination technique that requires the examinee to write on a piece of paper (or enter in a computer) a response to a question. The question might be anything from a multiple choice question to a true/false question to a question that asks for an essay analyzing the substantive and procedural law issues and the lawyering skills performance depicted in a videotaped vignette of a lawyering performance. This latter exam might include a client case file and it could be an in-class or a take-home exam. For a discussion of a variety of written test formats, see Steven Friedland, *Creating Effective Multiple-Choice Exam Questions* (unpublished paper presented to Gonzaga University School of Law, Institute for Law School Teaching, Conference on “The Science and Art of Law Teaching” (July 1995)) (on file with the author).

and with special notice, and I focus in this Article on such immediately recognizable instances of "examination."

In particular, I focus here on "paper and pencil" tests. For many practical and economic reasons, it is this type of test that predominates in most academic settings, including law schools.¹² On bar examinations in this country, written tests are the only tests. For the same kinds of resource reasons that written exams are the generally favored method throughout the law school curriculum, they may also be the only feasible way to expand skills testing. Individualized evaluation of a student's actual performance of oral lawyering skills requires an individual to observe the performance and give a grade of some sort. Clinicians do exactly this, and it is such personalized supervision and feedback that characterizes clinical education at its best. But when large numbers of potential examinees are involved—especially in bar exams or in most first year law school exams—such personalized evaluations by full-time faculty simply are not feasible. To offer the kinds of one-on-one observations by full-time faculty that take place in clinics to significantly larger numbers of students would necessitate seemingly prohibitive costs for increases in faculty. Essentially for these reasons, if we want to test lawyering skills for large numbers of students we must look at the possible "paper and pencil" methods of doing such evaluations.¹³

Part I of this paper first examines how lawyering skills are evaluated in the live-client clinic. Before we can consider whether written tests for lawyering skills make sense, it seems appropriate to look first at the one-on-one, personal methods of evaluating skills that in-house clinical teachers now use. What criteria are we using? How are we conducting the evaluations? Do our evaluations meet the require-

¹² Further, in law schools the written exam typically is a single test, given at the end of the semester or a course, that counts for all or most of the course grade. See Kissam, *supra* note 5, at 437.

¹³ There are, however, ways in which individualized critiques of student performances of oral lawyering skills could be provided without using *full-time* faculty, and perhaps without directly involving faculty at all. For example, adjuncts or even student teaching assistants could be used to provide feedback or even a grade on a student's skills performance test. See Jay Feinman, *Teaching Assistants*, 41 J. LEGAL EDUC. 288 (1991). In medical schools, trained actors are being used both to roleplay a patient in a simulated doctor-patient setting and to evaluate the performance of the student-doctor. See, e.g., Howard Barrows, *An Overview of the Uses of Standardized Patients for Teaching and Evaluating Clinical Skills*, 68 ACAD. MED. 443 (1993). While there is much that can be said in support of evaluation/testing methods that do not require full-time faculty, that discussion is a subject for another paper. (I am now conducting an experiment in adapting the medical school trained-actor technique to legal education, and expect to complete a paper describing those results in the near future.) Useful as such techniques may be, in any case, they are unlikely to eliminate the need for skills tests, especially on the bar exams—provided that, as this Article seeks to demonstrate, effective skills examinations can be designed.

ments of validity and reliability that are often used to assess testing fairness? Should we be using written tests to replace or supplement our individualized oral or written critiques, and if not, why not? I ultimately conclude that although one-on-one evaluations are *not* perfect, they certainly can be effective and fair. If pencil and paper tests have a role in clinics at all, it will likely be a limited one.

The situation is very different outside the clinics. In Part II, I move on to the non-clinic context, and describe the other settings in which lawyering skills may now need to be evaluated. There are two such contexts that are especially important. First, in law schools, these skills are now being taught in quite traditional classes, often large traditional classes. Certainly skills such as interviewing, counseling, negotiation, fact investigation and trial advocacy are likely to be taught at much lower levels of intensity in these classroom settings than in a clinic. But they can and are being taught there, and so they need to be evaluated there as well. Second, after law school, such skills are now increasingly being considered as potential subjects for the bar exam. The prospect of this modification of the bar exam offers the very welcome possibility of making the bar a somewhat more meaningful test of the ability of future lawyers. Neither in large classes nor in the bar exam, however, is anything like individualized observation of cumulative lawyering performance by full-time faculty going to be possible.

Instead, the usual testing method in both these settings is the written exam, and, as I will show, these tests have typically not been designed to offer any meaningful evaluation of students' mastery of most lawyering skills. What remains to be seen is whether written exams could be designed that would do a better job. This is a question that clinicians should help answer. Just as we have experimented with and developed new ways to integrate skills teaching into more traditional law school contexts,¹⁴ we should now address the question that these efforts have generated—that is, the issue of whether and how we might move toward evaluating or testing these skills in non-clinical settings.¹⁵

Addressing this issue is the subject of Part III. Here I examine what these paper and pencil alternatives for skills testing are, their possible utility and their limitations. In particular, I describe and as-

¹⁴ See, e.g., Lawrence M. Grosberg, *The Buffalo Creek Disaster: An Effective Supplement to a Conventional Civil Procedure Course*, 37 J. LEGAL EDUC. 378 (1987); Philip G. Schrag, *The Serpent Strikes: Simulation in a Large First-Year Course*, 40 J. LEGAL EDUC. 555 (1990).

¹⁵ Dean Harbaugh certainly has taken the lead in this area. See Joseph D. Harbaugh, *Examining Lawyers' Skills*, BAR EXAMINER, Nov. 1990, at 9.

sess the performance test (in which the examinee receives a simulated lawyer's file, from which he or she must ferret out the relevant facts and law in order to prepare some kind of a lawyering document such as an opinion letter); the videotaped performance test, in which the examinee watches a videotaped lawyering vignette and then evaluates the lawyer's performance shown on the tape; and interactive videotaped exams, which could even enable students to participate—via computer—in the performance of a lawyering task. I conclude that these methods, though not likely to provide as profound a measure of students' lawyering skills as a full-time clinical teacher can achieve with his or her students, *are* capable of providing meaningful forms of evaluation of a substantial range of skills.

I. TEACHING AND EVALUATING LAWYERING SKILLS IN THE CLINIC

Observation followed by feedback and evaluation is the clinical education equivalent of the traditional end-of-the-semester exam. This seems to me to be irrefutable. We observe and then comment, criticize, or cajole (and sometimes we intervene directly because a performance is deficient and a client's interests are at stake); we evaluate; and ultimately we communicate our evaluation to the student, whether with a grade or through a written or oral evaluation.¹⁶ This evaluation process is a cooperative and collegial one and we try to conduct it in a supportive and pleasant manner. But it still is a test, or at least our equivalent of one. The basic question is: how are we doing in our testing role? To answer this question we must first consider whether lawyering skills can truly be evaluated at all, and then focus on whether the particular methods of evaluation favored by in-house clinical teachers meet the criteria of validity and reliability by which tests are judged.

A. *Can Lawyering Skills Be Evaluated?*

A fundamental premise of all clinical skills teaching is that there are generally accepted principles of competent skills performance.¹⁷

¹⁶ The formal grade—whether a letter grade or a “pass”—also communicates our evaluation to outsiders, such as prospective employers.

¹⁷ There are numerous compilations of the central elements of the main interpersonal lawyering skills, including AMERICAN BAR ASSOCIATION SECTION OF LEGAL EDUCATION AND ADMISSIONS TO THE BAR, TASK FORCE ON LAW SCHOOLS AND THE PROFESSION: NARROWING THE GAP, LEGAL EDUCATION AND PROFESSIONAL DEVELOPMENT—AN EDUCATIONAL CONTINUUM (1992) [hereinafter MACCRATE REPORT]; AMERICAN LAW INSTITUTE-AMERICAN BAR ASSOCIATION COMMITTEE ON CONTINUING PROFESSIONAL EDUCATION, SKILLS AND ETHICS IN THE PRACTICE OF LAW (1993); AMERICAN LAW INSTITUTE-AMERICAN BAR ASSOCIATION COMMITTEE ON CONTINUING PROFESSIONAL EDUCATION, A PRACTICAL GUIDE TO ACHIEVING EXCELLENCE IN THE PRACTICE OF LAW:

Reading and studying these principles and the interpersonal theories underlying them can contribute to students' understanding of what constitutes a good or bad performance. In fact, reading about these skills, typically in one of the leading clinical textbooks,¹⁸ is an important part of most clinical courses. Certainly there is much more that is involved in the students' learning process: they also observe teacher modeling of skills performance, view video vignettes of good and bad performances, do simulation exercises, and give critiques of others' performances. Above all, the students receive meaningful personal feedback on their performance, typically through one-on-one tutorial sessions. But in all these other contexts, the skills theories identified in their reading provide the framework for critique and analysis.

Some might argue that when it comes to the interpersonal aspects of many lawyering skills (counseling or negotiating, for example), the factors of judgment, personality and strategic decisionmaking are so overwhelmingly subjective that they simply cannot be evaluated in anything even approaching an objective fashion.¹⁹ If this were so, then clinicians trying to critique their students' work in these respects would be reduced to saying "I know a good interview when I see one," or "that was a horrible cross-examination."

I believe that twenty-five years of clinical education have demonstrated that "feedback" of this sort is unhelpful and that clinicians can do better than this. The principles of competent skills performance to which I referred earlier do not remove all subjectivity from

STANDARDS, METHODS, AND SELF EVALUATION (1992) [hereinafter ALI-ABA, PRACTICAL GUIDE].

For further reflection on the nature of lawyering competency, see generally AMERICAN BAR ASSOCIATION SECTION OF LEGAL EDUCATION AND ADMISSIONS TO THE BAR, TASK FORCE ON LEGAL COMPETENCY, REPORT AND RECOMMENDATIONS OF THE TASK FORCE ON LAWYER COMPETENCY: THE ROLE OF LAW SCHOOLS (1979); H. Russel Cort & Jack L. Sammons, *The Search for 'Good Lawyering': A Concept and Model of Lawyering Competencies*, 29 CLEV. ST. L. REV. 397 (1980); Gordon E. Gee & Donald W. Jackson, *Bridging the Gap: Legal Education and Lawyer Competency*, 1977 B.Y.U. L. REV. 695; Wallace Loh, *Introduction: The MacCrate Report - Heuristic or Prescriptive*, 69 WASH. L. REV. 505, 509 (1994); *Symposium on Legal Education*, 30 N.Y.L. SCH. L. REV. (1985).

¹⁸ See, e.g., ROBERT M. BASTRESS & JOSEPH D. HARBAUGH, *INTERVIEWING, COUNSELING, AND NEGOTIATING: SKILLS FOR EFFECTIVE REPRESENTATION* (1990); DAVID A. BINDER, PAUL BERGMAN & SUSAN C. PRICE, *LAWYERS AS COUNSELORS: A CLIENT-CENTERED APPROACH* (1991); DONALD G. GIFFORD, *LEGAL NEGOTIATION: THEORY AND APPLICATIONS* (1989).

¹⁹ For the view that "we lack any objective measures of competence in practice," see Daniel J. Givelber, Brook K. Baker, John McDevitt & Robyn Milano, *Learning Through Work: An Empirical Study of Legal Internship*, 45 J. LEGAL EDUC. 1, 21 (1995). See also Jonathan Rose, *The MacCrate Report's Restatement of Legal Education: The Need for Reflection and Horse Sense*, 44 J. LEGAL EDUC. 548 (1994) (criticizing the MACCRATE REPORT, *supra* note 17, for developing a formalized list of lawyering skills without assigning weights to each skill reflecting their relative importance).

performance or evaluation—nor do the criteria for effective answers to traditional doctrinal essay questions—but they do rest on careful analytical breakdowns of the component parts needed to achieve effective performance. The comprehensive taxonomy of skills and values set forth in the MacCrate Report reflects exactly this sort of analysis of lawyers' work.²⁰ It seems fair to say that clinicians agree, for the most part, on the central tenets of most lawyering skills, ranging from courtroom trial advocacy to mediating a dispute.²¹

If there are effective and ineffective ways to perform these skills so that students can be taught, it seems axiomatic that the same criteria used for teaching can be used to evaluate a performance, whether that performance is written or oral. Evaluating an instance of lawyering performance, especially an oral performance, usually requires different methods of analytical assessment than would be used in evaluating a written essay answer, but this is a difference of degree and not of kind. In principle, there is no reason why lawyering skills should not be evaluated—and making such evaluations is central to training future lawyers.

B. How Do Our One-On-One Evaluations Fare When Subjected to Standards of Testing Fairness?

The question that remains is whether the methods of evaluation used in clinics are fair ones. The two primary criteria of testing fairness are content validity and reliability.²² These criteria should govern whether the testing method is the grading of a written exam or the oral critique of a lawyering skills performance, and whether the exam is a "high stakes" examination such as the bar exam or a test that covers much less material and is much less significant in its impact on the examinee. In this section I assess the extent to which the clinical one-on-one evaluation does meet these standards.

1. Validity

For a test to be valid, it must satisfy two requirements.²³ First, it

²⁰ See MACCRATE REPORT, *supra* note 17, at 135-221. Cf. Anthony G. Amsterdam, *Clinical Legal Education—A 21st Century Perspective*, 35 J. LEGAL EDUC. 612 (1984) (also analyzing the component elements of effective lawyering).

²¹ Just how complete this consensus is may well be a matter of debate. It would be a mistake, however, to deny the possibility of fair grading so long as *any* disagreement exists, for complete unanimity on this or any other legal subject will always be hard to attain, and I believe that the level of consensus is substantial.

²² See generally ROBERT L. EBEL & DAVID A. FRISBIE, *ESSENTIALS OF EDUCATIONAL MEASUREMENT* (4th ed. 1991); Julia C. Lenel, *Test Validation: What It Is and How It Should Be Done*, BAR EXAMINER, Aug. 1991, at 5; Douglas Kurdys, *Grading Essay Answers: The Issue of Reliability in Essay Scoring*, BAR EXAMINER, Nov. 1990, at 22.

²³ "Validity" appears to be a term of somewhat fluid meaning. I follow the explanation

must be an assessment of a skill or knowledge that is relevant to the overall purpose of the test. On a physician's licensing exam, for example, a test of one's ability to cook would be irrelevant and therefore invalid. Second, the test must call for test answers or responses that will permit an inference that the knowledge or skill being tested has been demonstrated. Multiple choice questions calling for doctrinal knowledge of real property law, for instance, would permit an inference that a correct answer denotes knowledge of real property law, but obviously would not permit such an inference about patent law.²⁴

Clinical educators' one-on-one observations and evaluations of their students should be highly valid. The skills being evaluated in typical clinics are certainly relevant to assessing students' competence as lawyers. Indeed, there would seem to be a strong consensus that such skills as interviewing or trial advocacy are essential for competent lawyering. Similarly, there should be no doubt that inferences about student skills are permissible based on direct observation of a student's lawyering skills performance, provided that the evaluator used appropriate criteria for the assessment.²⁵ Since, as we have already seen, there is substantial agreement among clinicians on the components of effective lawyering, this aspect of validity should be assured.

To put the matter plainly, an evaluation of a student's skills performance in a real case is the most direct and valid way in which to evaluate the student's ability and skill in doing what is observed.²⁶ In a simulation exercise, the inferences permissible as to the student's ultimate ability to perform the skill are not as strong—the absence of the real life context is a significant distinction—but observation of such exercises should still provide a quite valid test of the skill in question.

of the concept in Lenel, *supra* note 22, at 5-7. See also Marcia A. Kuechenmeister, *A Performance Test of Lawyering Skills: A Study of Content Validity*, BAR EXAMINER, May 1995, at 23.

²⁴ How one defines the purpose of a test becomes crucial. For example, if the purpose of a bar exam is to test knowledge in 15 bodies of doctrinal law, the "validity" of the exam is going to be quite different than if the purpose of the exam is to test an applicant's minimum competency to practice law, for competency requires much more than doctrinal knowledge.

²⁵ The student should also know what the criteria are, so that he or she can prepare as well as possible to satisfy them. Thus in any clinical context, whether an actual case or a simulation exercise, an important element of fairness would be the advance distribution to students of the criteria to be used for the assessment. The same would be true in a doctrinal course; the real property student cannot reasonably expect to be tested on patent law.

²⁶ As Jill Harker has commented, "[a]n example of the ultimate performance test would be a group of experts watching an examinee in the practice of law for a reasonable length of time and then deciding whether the examinee is competent to be admitted to the bar." Jill K. Harker, *Examining Examinations* (Jan. 1992) (unpublished paper presented at AALS 1992 Annual Meeting, San Antonio) (on file with the author).

2. Reliability

If different teachers, ostensibly using the same criteria, would grade the same student differently for the same performance, then the grading process is not consistent, and therefore not "reliable."²⁷ Is clinical evaluation based on a one-on-one observation of the student's performance of the lawyering skill a reliable system? Here, the clinical method is more vulnerable to criticism. The subjective bias of a teacher is difficult to prevent. The students whom the teacher likes may also become the ones who receive the most positive evaluations, despite the teacher's best efforts to avoid such unfairness.

Other points have been made as well as to why clinicians' grading may not be reliable. First, there is the very real concern that when a clinical teacher is integrally involved in the representation of a real client, that responsibility may distort or interfere with the ability to evaluate and grade the student.²⁸ Assuring that the client's legal interests are protected is not, after all, necessarily the same thing as evaluating the quality of a student's performance—even a good student may get in over her head. If the clinician cannot clearly, and unemotionally, focus on the quality of the skills performance(s) on a real case, the evaluation method of direct observation leaves something to be desired. Second, there is a danger that clinicians will invest so heavily in guiding their students' work that when they evaluate their students' performance they are, in effect, evaluating their own lawyering. If this is so, then our confidence in the reliability of their evaluations would certainly be diminished.²⁹

Clinicians have used a variety of possible antidotes to minimize potential grader prejudice, and thus increase the level of reliability. For example, if two teachers are in the clinic, each can independently

²⁷ Professor Friedland, borrowing from Professor Paul Wangerin, gives the following very useful explanation of "reliability" errors, which he describes as those "attributable to inconsistency in grading. To illustrate reliability errors, he [Wangerin] uses the example of an old watch that runs faster on some days than others. . . . Thus, an inconsistent measure is one that changes over time or from paper to paper.

"Professionals rank the level of reliability of a testing instrument on a scale of zero to one. Tests that are highly reliable and produce absolutely no error are given the rating of a perfect 1.0; tests that are no more reliable than chance are given the lowest reliability rating of 0.0." Friedland, *supra* note 11, at 7 (citing Wangerin, *supra* note 5, at 24-25).

The "professionals" to whom Friedland refers are the psychometricians, social scientists who use statistics to design tests and to analyze all of the variables relevant to the relationship between the results of a test and the test design to ensure fairness and objectivity.

²⁸ See James H. Stark, *Preliminary Reflections on the Establishment of a Mediation Clinic*, 2 CLIN. L. REV. 457, 497-501 (1996). Many, perhaps most, clinics do not require faculty to become so deeply involved in the actual handling of the cases—but some do.

²⁹ See Robert J. Condlin, *Socrates' New Clothes: Substituting Persuasion for Learning in Clinical Practice Instruction*, 40 MD. L. REV. 223, 233-38, 267-70 (1980).

grade a student, presumably using the same criteria. The student may also be asked to complete a self-evaluation, which the teachers use as a check on their own judgments. Faculty can also obtain evaluations from third parties who participate in or observe the lawyering performance.³⁰ Consistency is the objective. We should not take the reliability of our grading for granted, and if we find that grade disparities exist, some corrections are necessary.

C. *Should Clinicians Use Written Skills Tests?*

If we have valid criteria for a competent skills performance and if we have taken effective steps to ensure consistency and reliability in our use of those criteria, then our observational method of evaluating interpersonal lawyering skills would be optimal. In this context, with faculty providing one or more such evaluations for each clinic student, a separate end-of-the-semester cumulative exam would seem superfluous, and could be counterproductive. In a live-client clinic, after all, such a cumulative performance exam probably could only be done in a simulation exercise, whereas some of the evaluations earlier in the semester would likely have involved student work with real clients. The simulation could not replicate the immediacy element in the live-client activities. Moreover, the students could resent an examination based on a simulation, precisely because it might not reflect their ability to handle real-life exigencies. In addition, the collegial atmosphere and the acceptance of open, constructive critiques that most of us encourage in our clinics could be adversely affected if a formal test—"just like the rest of law school"—were employed in the clinic.

At the same time, it may be that we should be using, or at least considering the use of, supplemental methods of evaluating our clinic students. I see at least two reasons for this conclusion. First, if there is any accuracy to the "reliability" criticisms that I have described, then use of a more "objective" method of evaluation could be an appropriate and useful complement to one-on-one observations. Second, work on real cases often is unpredictable and uneven, and some students may be deprived of the chance to perform many, even most, of the skills being taught in the clinic. A graded simulation exercise may be useful for either of these reasons, even if it is not a complete substitute for observing a performance with a real client. Alternatively, it may even be appropriate for clinicians to consider some kind of a

³⁰ In one simulation course I teach, outside lawyers participate in a negotiation exercise and are asked to provide some feedback using a standardized form. As noted above, *see* note 13 *supra*, part-time adjunct teachers or role-playing actors may also be used to evaluate student performances.

written test, of the sort discussed below.³¹

II. TEACHING AND EVALUATING SKILLS IN A NON-CLINIC CONTEXT

The current reality is that individualized clinical feedback is a relatively rare experience for law students. Financial constraints on law schools have severely limited the number of clinical opportunities. Even students who do enroll in a clinic typically take only a single clinical course. We may seriously dispute law school claims that financial pressures prevent any expansion of clinics. In the meantime, however, the reality is that clinic spaces are limited.

We thus are faced with the question of whether the in-house clinics are the only place where we can teach or evaluate students' interpersonal lawyering skills. The answer, I believe, is "no", and I will begin this Part by describing the range of sites, in law school and on the bar exam, where such skills are now, or could be, taught or assessed. But I will go on to explain why it is also clear that the traditional mode of evaluation in non-clinical contexts—the standard law school exam—is poorly suited to the task of evaluating these skills. Then it will be time, in Part III, to describe evaluation methods that educators and evaluators can employ to measure skills in these new contexts.

A. *In What Contexts, Apart from the Clinic, Can We Teach and Evaluate a Broad Range of Lawyering Skills?*

We need to consider two very different settings—the law schools and the bar exams. Clearly, the law schools have responded over the last decade to the challenge of providing more ways to teach lawyering skills. Many new skills offerings have been established to complement the live-client clinic. These include pure simulation courses and introductory lawyering courses, as well as more traditional doctri-

³¹ I am not aware of any comprehensive data that address the issue of whether graded simulations or written tests are used by in-house clinicians as a supplement to casework evaluations in determining students' grades. My impression is that some in-house clinicians do use graded simulations, but few, if any, use written tests. (Robert Seibel and Linda Morton have found in their survey of externship courses that only one of the 98 programs they surveyed gives grades based on a written exam. Robert F. Seibel & Linda H. Morton, *Field Placement Programs: Practices, Problems and Possibilities*, 2 CLIN. L. REV. 413, 435 (1996).) Pure simulation courses are a different matter. In my four-credit, upper class "Negotiating, Counseling and Interviewing" course, for example, I use both graded simulation exercises and a final exam as components of the students' final grade.

See also Roy T. Stuckey, *Results of Survey on Testing in Clinical Courses and News About the NCBE's Performance Test*, SECTION ON CLINICAL EDUCATION NEWSLETTER (Dec. 1993) at 9 (in this survey, to which very few teachers responded, a handful of clinical teachers indicated they were using tests of various sorts).

nal courses which have incorporated various kinds of simulation skills exercises into their work. Even externship, workshop or field placement courses approach the teaching and evaluation of skills in somewhat different ways than are characteristic of the live-client clinic.³² Yet in all of these courses, skills may be taught, and should be evaluated.

Change in the bar exam, by contrast, has been considerably slower. Even now, state bar exams do not test applicants for their ability to perform all or even most lawyering skills necessary to competent lawyering.³³ To the contrary, the typical bar exam tests doctrinal knowledge (of varying quantities of law) and a very limited number of skills, usually including legal analysis and essay writing. As we will see, however, there is strong reason in principle to believe that the bar exam should be reshaped to cover a wider range of these skills.³⁴

1. *The Law School Classroom*

Expansion of the scope of lawyering skills being taught or used in the classroom is occurring on several fronts. As this brief survey will reflect, skills training now takes place in settings ranging from doctrinal courses taught in classic large classes to much more novel classes centered around skills issues.

Large-classroom skills education includes well-established doctrinal courses in which some elements of applied lawyering skills are presented or taught to students. In civil procedure, for example, an

³² Placement courses share with in-house clinics a focus on the student's experience in actual lawyering roles, and a concern with ensuring that that experience receives appropriate evaluation. But the modes of evaluation in the two contexts (in both of which I teach) are not the same. The main difference between placement courses and in-house clinics in this regard is that externship teachers typically are not in a position to evaluate student performance in actual cases; such evaluation must come, if at all, from the field placement supervisor rather than the full-time faculty member. Similarly, even if there is an on-campus seminar component, it may or may not include observational feedback by the externship teacher. Again, it is only in the clinic where direct observation typically is the norm. But in some clinics, as in externships, there may be situations in which no faculty member accompanies the student or views tapes of the student in action. In those clinics, supplemental evaluation/testing methods may become especially desirable.

³³ Even the three states (Alaska, California and Colorado) that now include a performance test as a regular part of their bar exams still do not test for most lawyering skills. On a positive note, however, the National Conference of Bar Examiners now plans to offer a Multistate Performance Test beginning in February 1997. See text at notes 57-80 *infra*.

³⁴ See Committee on Legal Education and Admission to the Bar, Association of the Bar of the City of New York, *Report on Admission to the Bar in New York in the Twenty-First Century—A Blueprint for Reform*, 47 RECORD OF THE ASSOCIATION OF THE BAR OF THE CITY OF NEW YORK 464, 507 (1992) [hereinafter CITY BAR REPORT]. Anthony E. Davis was the chair of the Committee and I was the Reporter for this Committee Report, one of whose primary criticisms of the bar exam was for its failure to test for proficiency in the range of skills necessary for minimally competent lawyering.

increasing number of teachers are incorporating one or another lawyering activity beyond the traditional analysis of appellate opinions.³⁵ The students might be asked to conduct a deposition or argue a motion or draft a complaint or set of interrogatories. Can or should students' experiences with such assignments be integrated into the testing processes in the course?

A second kind of classroom skills training is the first year "Introduction to Lawyering" course.³⁶ Sometimes such a course is integrated with first year writing,³⁷ at other schools it is a stand-alone course, often a large one.³⁸ As in the civil procedure example given above, the depth of students' immersion into lawyering skills may vary from very modest to relatively intense. In any case, however, the question remains the same: should the extent of classroom skills teaching be reflected in the testing process used in the course?

Still other law school classrooms in which skills are taught include simulation courses on such topics as trial advocacy; interviewing, counseling and negotiating; and alternative dispute resolution. Field placement courses with a seminar component also frequently emphasize skills education. In courses like these, a teacher's direct observation of simulated performances often constitutes a large part of any grade. In our trial advocacy course, for example, a student's performance in an end-of-the-semester simulated jury trial is the most significant part of the course grade. The question is whether written tests might be a useful supplement to such direct observation and evaluation. The answer, here as in the other classroom courses, is entirely dependent on the utility and practicality of possible methods of skills testing, the subject of Part III below.

2. *The Bar Examination*

Individualized evaluation of an examinee's oral lawyering skills is not now occurring on any bar exam, and is not likely to occur. Nor do most bar exams include any questions or problems calling for a written answer that would permit an evaluation of the examinee's competency in the vast majority of practical lawyering skills. Whether the bar exams should include some form of testing for such skills ulti-

³⁵ See, e.g., Grosberg, *supra* note 14; Schrag, *supra* note 14; *Symposium on Simulations*, *supra* note 4.

³⁶ Such courses are also sometimes named "Legal Method" or "Introduction to the Profession."

³⁷ For example, the "Lawyering" course at New York University School of Law is an eight-credit, year-long course that teaches writing, legal reasoning and oral lawyering skills.

³⁸ At New York Law School, where I teach, this course, "Lawyering," is taught in a large classroom (100-120 students) using video and live demonstrations, detailed case files and limited student roleplaying.

mately depends, of course, on the effectiveness and feasibility of such tests. But it should be clear that if effective tests can be designed, there would be good reason to use them in the bar exam.

The ostensible purpose of bar exams, after all, is to protect the public from incompetent lawyers.³⁹ This is not the place to debate the utility of bar exams.⁴⁰ Regardless of one's position on whether to retain bar exams, however, there is increasing and widespread agreement that "minimum competence" encompasses a wide range of skills,⁴¹ including many not now being tested on these exams. It also seems reasonably clear that there is no likelihood that bar exams will be eliminated in the near future.⁴² If the passing of a bar exam can be made to be a more accurate reflection of an applicant's minimum competence to practice law, and if this can be done without adding to the costs of bar exams and the agony of bar examinees, then surely it

³⁹ The "purpose [of the bar exam is] to protect the public." NATIONAL CONFERENCE OF BAR EXAMINERS, CODE OF RECOMMENDED STANDARDS FOR BAR EXAMINERS Standard 18 (1987), reprinted in NATIONAL CONFERENCE OF BAR EXAMINERS, THE BAR EXAMINERS' HANDBOOK 70:4 (Stuart Duhl ed., 3rd ed. 1991) [hereinafter BAR EXAMINERS' HANDBOOK]. For a similar appraisal, see George Stevens, *Diploma Privilege, Bar Examination or Open Admission*, 46 BAR EXAMINER 15, 21 (1977) (the reason for the adoption of the written bar examination requirement was "to protect the public from the harm done by the poorly prepared or incompetent judge or lawyer"). So, too, a recent study commissioned by the New York Court of Appeals described the bar examination as "a licensure examination" and said that, "as such, its purpose is to protect the public against incompetent lawyers." JASON MILLMAN ET AL., AN EVALUATION OF THE NEW YORK STATE BAR EXAMINATION (1993) [hereinafter NEW YORK BAR EXAM EVALUATION].

⁴⁰ Compare James P. White, *Where the Testing and Admission Process Should Be in the Year 2000*, BAR EXAMINER, Feb. 1985, at 35, 36 ("Bar Examinations are 'necessary and proper; they provide a stimulus to law schools, a means of encouraging the schools to do the best job they can in legal education.'") (quoting Erwin Griswold, *In Praise of Bar Examinations*, A.B.A. J., Jan. 1974, at 81) with CITY BAR REPORT, *supra* note 34 (vigorously criticizing the bar exam as an inadequate test of minimal competency and as a possibly racially discriminatory barrier to law practice).

⁴¹ See, e.g., JOHN P. HEINZ & EDWARD O. LAUMAN, CHICAGO LAWYERS (1982); FRANCIS ZEMANS & VICTOR ROSENBLUM, THE MAKING OF A PUBLIC PROFESSION (1981); Leonard L. Baird, *A Survey of the Relevance of Legal Training to Law School Graduates*, 29 J. LEGAL EDUC. 364 (1987); Robert A. Fairbanks, *The Failure of American Legal Education: A Recommendation for an Integrated Legal Education Program*, 12 TULSA L.J. 627 (1977); John O. Mudd & John W. LaTrielle, *Professional Competence: A Study of New Lawyers*, 49 MONT. L. REV. 11 (1988); Robert A.D. Schwartz, *The Relative Importance of Skills Used by Attorneys*, 3 GOLDEN GATE U. L. REV. 321 (1973).

⁴² The irony is that while few if any observers believe that passing the bar reflects competence to practice, and many question its utility altogether, there also is very little support for eliminating the bar exam. Citing statistics collected by others, Professor Deborah Rhode writes that "[t]here appears to be little enthusiasm for modifying the existing examination system. According to several 1980's American Bar Association polls, between 83 percent and 88 percent of surveyed lawyers believed that bar exams should continue, although 70 percent also agree that the exams don't adequately measure the ability to practice law." DEBORAH RHODE, PROFESSIONAL RESPONSIBILITY—ETHICS BY THE PERVASIVE METHOD 65 (1994).

is time for such changes to be made.

B. What Are the Limits of Traditional Written Tests?

"Traditional" written tests come in several varieties, and they test more than one aspect of students' knowledge and ability. But a brief review of conventional testing will confirm that the scope of these tests is nonetheless severely limited.

Both in law school and on bar exams, the traditional essay question remains the most commonly accepted and probably the most frequently used written test format. A typical question will provide a fact pattern and ask the examinee to write an "issue spotting" essay or a bench brief for the judge deciding the case.⁴³ In addition to issue spotting, two other "intellectual functions" tested on these exams are the identification of legal authorities and the analysis or application of law to the fact pattern.⁴⁴ All of these abilities are part of "legal reasoning," though they certainly do not encompass all of the thinking that a lawyer must do on an actual case. Other functions or capacities tested on these exams may include organization of answers, writing ability, speed,⁴⁵ and ability to deal effectively with surprise.⁴⁶

Other written testing methods have been more exclusively concerned with measuring doctrinal knowledge.⁴⁷ A variation of the essay is the short answer exam—a question that calls for anything from a phrase to a paragraph or two in response to a fact pattern or some other kind of question. Multiple choice or true-false questions (sometimes both referred to as "objective tests") are a significant part of the bar exam and are also used in varying degrees in law schools. Increasing efforts are being made to refine multiple choice exams to test legal analysis and reasoning.⁴⁸

If our mandate as legal educators is limited to teaching doctrine,

⁴³ See generally Kissam, *supra* note 5.

⁴⁴ *Id.* at 440-41.

⁴⁵ Gary Spencer, *Report Calls Test of Speed Key for Bar Exam*, N.Y.L.J., July 12, 1994, at 1 (quoting NEW YORK BAR EXAM EVALUATION, *supra* note 39).

⁴⁶ Kissam, *supra* note 5, at 453-57. In a similar vein, Erwin Griswold suggested that the bar exam tests the applicant's ability to withstand the stress inherent in law practice. Griswold, *supra* note 40, at 81, *quoted in* BAR EXAMINERS' HANDBOOK, *supra* note 39, at 10:5.

⁴⁷ The proposition that a major goal in law school and on bar exams is to teach and test on the broadest possible list of doctrinal subjects rests on dubious if not totally illusory assumptions. As Judge Harry Edwards has stated: "Law Schools should not seek to provide students a comprehensive knowledge of legal doctrine, for it simply cannot be done." Harry Edwards, *The Growing Disjunction Between Legal Education and the Legal Profession*, 91 MICH. L. REV. 34, 57 (1992). A critical threshold issue when considering "reform" of the bar exam is how much doctrinal law should be covered in a bar exam, and a similar question exists with regard to examinations in law school. These important questions, however, are beyond the scope of this Article.

⁴⁸ See generally Friedland, *supra* note 11.

writing, and the reasoning and analysis required to write about doctrine, it seems needless to consider other testing methods. But if we have a broader calling, as a recognition that competent lawyering entails much more than these abilities suggests, then we should consider additional assessment techniques. For what the issue spotting essay question and the objective or short answer questions of traditional exams assuredly cannot test for is the fuller array of applied lawyering skills.⁴⁹

These traditional written exams do not assess the skills of interviewing or counseling a client, negotiating with an adversary, or cross-examining a witness. They do not address the important skill of problem-solving. Nor do they typically call for the drafting of various kinds of documents that lawyers are called upon to prepare, such as motions or opinion letters or statutes. Most fundamentally, these tests do not examine a student's (or a bar exam applicant's) ability to ferret out key facts from an undigested body of data or to distinguish relevant from irrelevant facts. These are significant gaps in the scope of an exam, if the goal is to measure law students' acquisition of practice skills or to assure the public that licensed lawyers are competent to practice. The question then is whether there are ways to address these shortcomings.

III. ALTERNATIVE WRITTEN TESTS OF A BROADER RANGE OF LAWYERING SKILLS

At least three different kinds of paper and pencil exams have been or could be used to test a range of skills beyond the scope of what traditional tests can measure. The first is the "performance test" (PT), a testing technique that has been used as a regular part of the bar exam in California (for more than a decade) and in two other states.⁵⁰ This is a question that calls for the examinee to review a file that contains both factual material and law and then perform a task similar to what a lawyer might be asked to complete—to draft a negotiation plan, for example, or prepare a witness examination outline.

The second possibility is what I call a video PT. In addition to the PT file, this form of exam includes a video vignette that depicts some aspect of the legal and lawyering challenge in the problem. The ex-

⁴⁹ A recent exhaustive study of the New York State bar examination broadly affirmed the value of this exam, but also concluded: "We are persuaded that the present examination does not test many skills deemed important to be a successful and competent lawyer. . . . We approve measuring a broader range of skills as well" *NEW YORK BAR EXAM EVALUATION*, *supra* note 39, at 3-10 to 3-11.

⁵⁰ I have served as a member of a drafting team responsible for preparing a question for the California PT test.

aminee then is asked both to complete a PT lawyering task of the sort outlined above and to provide a critique of the lawyer's performance that the examinee has viewed on the tape.

The third paper and pencil model is an interactive computerized video exam. (Here, of course, the "pencil" has been replaced by the computer keyboard.) Like the "video PT," this interactive exam could also include a PT case file that would be the basis for the events shown on video. Such an exam might simply ask the user to answer various questions about what is shown at particular points in the videotape, which would stop while the examinee gave his or her answers. But it might also go further by including multiple alternative courses of action that an examinee might pursue by choosing a particular answer.

Three overriding issues are central to any consideration of these written test possibilities. First, does the fact that someone can write a cogent analysis of how to perform a skill or a critique of another's performance mean that the person can actually perform the skill? Second, is each of these tests "valid and reliable" enough to be a reasonable substitute, or even a supplement, for the clinician's one-on-one tutorial critiques? Third, how feasible would it be to implement each technique, either in the classroom or on a bar exam? I will take up these questions in this Part; I will conclude that the PT itself is already demonstrably a valuable form of testing, and that each of the other two forms of skills examination also offers real promise of adding to our complement of tools for evaluating skills.

A. Performance Test ("PT")

1. What Is the PT and What Does It Test?

The PT is a written question that asks the test taker to apply the law in context, in a fashion similar to the way a lawyer might be called on to perform in a real case.⁵¹ Several elements contribute to this goal. First, rather than give the examinee an abbreviated and distilled version of the relevant facts, the PT generally provides an undigested collection of facts. The "File" might include deposition transcripts, original contracts, correspondence, and similar materials. By including

⁵¹ The term "PT" is somewhat misleading because this exam calls for only a "written performance" by the test taker. More colloquially, "performance" would connote both oral and written performance. Indeed, some might think a performance test would examine only oral performance. Nevertheless, while a more accurate name for this test would be "written PT," PT is the term used in the literature, and is the term I will use. For a comprehensive description of the performance test, see *Performance Testing: A Valuable New Dimension or a Waste of Time and Money*, BAR EXAMINER, Nov. 1983, at 12, 13 (comments of Armando Menocal) (edited transcript of an ABA panel discussion, held Aug. 1, 1983, concerning the California PT test).

such "real" law documents, the PT aims to present the complexity of reality. Like an actual lawyer, the examinee must now discern the relevant facts—facts, for example, that could be transformed into admissible evidence. Second, the law is usually provided in the form of actual statutes, administrative regulations, or case law in a "Library"; reading and analyzing this library, and applying it to the facts, requires the examinee to engage with the raw materials of the law, probably somewhat more than most traditional exams would. Third, the examinee must perform a lawyering task based on his or her evaluation of the facts and law. The specific requested tasks might range from preparing a counseling letter to a client, to drafting or analyzing a contract, to preparing a jury trial summation.

At a minimum, the PT test evaluates the skills of legal analysis and reasoning and written communication. In these respects, it achieves the same objectives as a traditional essay question.⁵² In addition, simply by asking the applicant to assess a factual file and to distinguish the relevant from the irrelevant, the PT partially tests fact investigation skills. The PT can highlight these skills even further by also asking the examinee to evaluate what if any deficiencies there were in the fact investigation and to propose additional or alternative fact-finding steps. By asking the test taker to integrate facts and law and prepare a written plan for carrying out some additional task, the PT can also evaluate to a limited extent the student's ability to perform that task. The student's ability to outline a direct examination of a trial witness, for example, reflects to some extent his or her skill in trial advocacy. Similarly, if the task is to draft a counseling outline, the exam will allow a partial evaluation, if not of the student's counseling abilities, at least of his or her knowledge of what is required for competent counseling. Some PT's have also tested applicants on ethics and tactical issues.⁵³

Writing about a particular oral lawyering skill clearly is not the same as performing the task. This is a distinct limitation in what is achievable in PT's.⁵⁴ Certainly talking or writing about counseling,

⁵² In some ways, one might consider a PT a limited open book test. See generally Jane Peterson Smith, *The July 1993 Performance Test Project*, BAR EXAMINER, May 1995, at 36; Committee of Bar Examiners of the State Bar of California, *Information Regarding Performance Tests*, March 1994 (unpublished document, on file with the author).

⁵³ See Jane Peterson Smith, *Performance Testing in California, 1983-89*, BAR EXAMINER, Aug. 1989, at 17.

⁵⁴ See Stuckey, *supra* note 7, at 4. Professor Stuckey correctly points out that while the PT is an important new development, it does not now, and probably never will, assess such important lawyering attributes as "integrity, conscientiousness, motivation or the ability to instill confidence. . . . Clinical methodology [i.e., direct observation of lawyering actions] is a diagnostic and evaluative tool as well as a method of instruction. It is a true performance test." *Id.* at 6.

for example, is different from actually counseling a real client, or even a simulated client. Indeed, some would argue that the ability to describe an activity has nothing to do with the ability to perform the activity.⁵⁵ But an astute written analysis does demonstrate an awareness and an understanding of the theories and problems underlying the performance of skills. Ultimately, such knowledge and understanding should contribute to the efficacy of the lawyer's actual performance of such tasks. Even pure book knowledge can contribute to one's lawyering skills, and what is required for the PT is more than that, for the student must actually apply that knowledge, at least on paper.

The value of such intellectual understanding of skills challenges may become more apparent if we imagine a survey course on alternative dispute resolution (ADR), a course without a clinical or even a simulation component. Even though such a course does not entail students' performance of the skills involved, the students do read, study and discuss both the underlying theories and various problems in the actual use of ADR. They gain a greater understanding of the mechanics of various ADR methods, the history of societal uses of ADR, the political responses to the ADR movement and various conceptual arguments about why ADR may or may not be fair or wise in particular situations. That knowledge and increased understanding simply must constitute at least a building block in the student's actual acquisition of ADR skills, even if—as is surely the case—other building blocks must also be in place before the student can become proficient in this work.

Similarly, the use of written texts in clinical and lawyering skills courses does not mean that a student's verbal or written mastery of the contents of such books is the equivalent of the student's ability to perform the skills being studied. But such textual and theoretical guidance certainly cannot hurt a student's simultaneous quest to learn "how to do it." To conclude otherwise would seem to return us once

⁵⁵ Professor Donald Schön, in discussing and writing about his concept of the "reflective practitioner," has explicitly observed "that doing something and answering a question about what you are doing are completely separate activities." Donald A. Schön, *Educating the Reflective Legal Practitioner*, 2 CLIN. L. REV. 231, 243 (1995). At the same time, Professor Schön very much believes in the value, if not the necessity, of reflecting through critical analysis—a "research process"—on one's performance of professional skills. It is through this analysis of the action that improvement can occur the next time around. *Id.* Thus, while writing about how to do something is not the same as doing it, I believe Professor Schön would agree that a written analysis of what to do in a certain situation, or a written appraisal of someone else's response to such a situation, could be helpful in educating and preparing the practitioner to perform the skills in question. To use his terminology, it is this post-event "reflection-on-action" that can lead to greater skills development, improved "reflection-in-action" and greater skills "artistry." *Id.* at 247.

again to the idea that skills cannot be analyzed, and to the notion that students can only learn them by entering practice and being required to "sink or swim."

Finally, it is important to emphasize that the PT generally asks the examinee to perform the requested lawyering task in a quite realistic context. Rather than directing students to reason from a one-paragraph hypothetical, the PT provides materials that are richer and more nuanced. These materials more closely approximate what a lawyer might have to delve into to solve a client's problem than a traditional exam normally would.⁵⁶ The student "performs" in writing, but that performance responds to a problem structured to present some of the complexity of real lawyering. Intuitively, it seems hard to deny that the student who does well on a PT is more closely replicating a wider range of the lawyering skills central to the notion of minimum lawyering competence than a student who writes a good answer to a traditional essay question.

2. *Is the PT a Valid and Reliable Test of Lawyering Skills?*

A recent study commissioned by the National Conference of Bar Examiners (NCBE)⁵⁷ affirmatively answered this question.⁵⁸ The purpose of this study (the "ACT Study") was to continue the NCBE's inquiry into the possible expansion of the use of PT's on bar exams. It analyzed the use of the PT in six states, three in which PT's had been included on bar exams for a number of years (California, Alaska and Colorado) and three that had not previously used PT's but agreed to participate in the study. In the latter three states (Georgia, New Mexico and Virginia), the applicants were given the option of taking the PT, on the condition that their answers could help, but not hurt their chances of passing. A sophisticated psychometric examination, the ACT Study analyzed both the demographics of the test takers and their comparative test scores on the PT and non-PT parts of the exam.⁵⁹ It also included a survey seeking responses of the applicants

⁵⁶ Surveys of PT test takers have demonstrated that practicing attorneys who take PT exams do better than novices, a result that confirms (or at least suggests) that the PT measures skills relevant and important to lawyers. See Harbaugh, *supra* note 15, at 14.

⁵⁷ The NCBE is a private non-profit organization that is affiliated with the ABA. For a critical analysis of the relationship of the ABA and the NCBE, see Roy T. Stuckey, *Preparing Lawyers for Law Practice: New Roles for the NCBE and the ABA*, BAR EXAMINER, May 1990, at 12.

⁵⁸ For the study results, see AMERICAN COLLEGE TESTING, RESULTS OF RESEARCH ON THE NCBE PERFORMANCE TEST PILOT ADMINISTRATION (1994) [hereinafter ACT STUDY] (report submitted to National Conference of Bar Examiners, Research & Development Committee, April 1994). See also Kuechenmeister, *supra* note 23, at 29.

⁵⁹ One complication in interpreting the PT exam scores was that some of the PT's were 90-minute tests, while others lasted 3 hours.

who took the exam.

While the number of skills tested on these July, 1993 PT's was limited (legal analysis, fact analysis, problem-solving and written communication), the results of the ACT Study were quite positive in several respects. In an independent analysis that the NCBE commissioned to review the data from this study and to assess the overall impact of the PT, Dr. Stephen Klein, perhaps the leading psychometrician analyzing bar exams, concluded that, "question-for-question, a PT task generally provides a more reliable score than does a typical essay question."⁶⁰

As to the validity of the PT, the study's outcome was also positive. The ACT Study concluded that "the technical quality of all [test] materials was rated 'Good' to 'Very Good' for all tasks"—a factor critical to the validity of the testing process—and "that the fundamental lawyering skills defined in the test specifications were indeed important."⁶¹ Dr. Klein concluded that the PT measures "an important ability . . . not fully measured by the other parts of the bar exam" and that "it appears that a PT would be a useful addition to most bar exams."⁶² Most importantly, the bar examinees themselves rated the PT as the best test [of the Multistate Bar Examination, the Multistate Essay Examination and the PT] in "measuring the ability to perform as an attorney."⁶³ Further, 55% said the PT should be continued as part

⁶⁰ Stephen P. Klein, Relationships Among MBE, Essay, and July 1993 Performance Test Scores at 2 (undated) (accompanying a Memorandum from Francis D. Morrissey, NCBE President, and Jane Peterson Smith, NCBE Director of Testing, to Interested Members of the Bar Admissions Community) (April 1994) (unpublished document, on file with the author). Dr. Klein explained that:

There are two relevant definitions of reliability for a bar exam. One definition refers to the degree to which a test rank orders examinees the same way regardless of the particular set of questions asked and in the case of an essay test, the attorneys who are assigned to grade the answers. The second and more important definition refers to the consistency with which different versions of the exam would make the same pass/fail decision about an applicant.

Id. Though his opinion is subject to a number of caveats, he in effect finds that the PT experiment satisfied, or that PT tests could satisfy, both of these definitions of reliability. For the similarly qualified, but still positive comments of the ACT Study itself on reliability, see ACT STUDY, *supra* note 58, at 23-25.

⁶¹ *Id.* at ii, iii. See Kuechenmeister, *supra* note 23, at 29 ("The skills outlined in the test specifications were judged by the panel [selected by ACT] to be essential in the practice of law . . .").

⁶² Klein, *supra* note 60, at 2. While Dr. Klein does not state what this "important ability" is, his finding seems to be a quite clear acknowledgement that the PT can meaningfully evaluate aspects of lawyering skills. This acknowledgement is particularly important since Dr. Klein had earlier been critical of the movement toward use of PT's on the bar exams. See note 66 *infra*.

⁶³ Charles S. Kuncie & Scott E. Arbet, *A Performance Test of Lawyering Skills: Candidate Perceptions*, BAR EXAMINER (May 1995), at 43, 44.

of the bar exam (25% were neutral and 20% were opposed).⁶⁴

Comparing the results that PT examinees achieve with those that the same people achieve on other parts of the bar exam is also illuminating. Psychometricians assert that those who do well on the short answer tests (the Multistate Bar Exam and the Multistate Professional Responsibility Exam) generally achieve the same results on bar essay exams and on the PT. This finding confirms the fairness of all of these measures.⁶⁵ It also raises the question whether the PT adds anything to the overall assessment of an examinee's competence. What is the point, then, of going to the trouble of preparing and administering the PT when the same people who do well on the PT also do well on the more easily prepared and administered tests?⁶⁶

The first answer to this question is that PT scores may *not* be identical to those the examinees obtain on other tests. The data from the ACT's six-state study are not clear, but they suggest that some test takers do better on PT's.⁶⁷ A close observer of PT's has also found from data on the California PT that as many as four percent of test takers do have better results and therefore benefit from the PT's. Another four percent, not surprisingly, do worse on the PT.⁶⁸ These figures suggest that almost a tenth of test-takers may perform differently on the PT than on other tests. If the PT test measures abilities the other tests do not directly capture, we would expect such differences.⁶⁹

Even if the correlation between PT scores and scores on other parts of the bar exam were perfect, however, there would still be good

⁶⁴ *Id.* at 47. Similarly, student evaluation forms and feedback in the courses in which PT's have been used at my law school have generally praised PT's as useful and fair exam techniques and recommended that they be retained.

⁶⁵ See Harbaugh, *supra* note 15, at 12; Klein, *supra* note 60, at 3; Smith, *supra* note 52, at 41.

⁶⁶ In criticizing a draft of the CITY BAR REPORT, *supra* note 34, in 1992, Dr. Klein then decried efforts at testing a wider range of lawyering skills because PT test takers generally fare the same on more traditional tests. Stephen P. Klein, Comments on the Committee's Discussion Draft, Mar. 2, 1992 (unpublished document included in Appendix H to the CITY BAR REPORT, *supra*, and on file with the author and at the Library of the Association of the Bar of the City of New York).

⁶⁷ "Attorneys who are already licensed to practice in a jurisdiction score statistically significantly higher on the PT than would be expected on the basis of their scores on the rest of the exam." Klein, *supra* note 60, at 3.

⁶⁸ Harbaugh, *supra* note 15, at 11. Citing studies of the California PT, Dean Harbaugh concludes that "there is some evidence that there are those who would benefit or be burdened by [the PT's] inclusion" in a bar exam. "[A]t the pass/fail decision point, the incorporation of the performance phase results in the passage of almost four percent of the total candidate pool who otherwise would fail, and the failure of another four percent of the applicants who, but for the performance test, would have passed."

⁶⁹ Moreover, we might find the differences becoming even greater as we expand the range of skills tested on the PT.

reason to administer the PT as part of the bar and as part of the examination in law school classes which incorporate skills education in their curricula. We would not, after all, have any reason to infer from this consistency of results that the PT was actually an *invalid* test. Nor would we accept the fact that there is a strong correlation between results on the LSAT, first year law school grades and bar exam results as a reason for saving time, energy and expense by admitting people to the bar based on their LSAT scores. (Similarly, if scores on a math test proved to correlate well with bar exam results, we would not be likely to install the math test as a simplified substitute for the more cumbersome bar exam.)

While the LSAT or even a math exam might have predictive value, passage of such tests would in no way reflect the knowledge or the skills necessary to practice law. The most important point about both the ACT Study and the earlier California study was that they demonstrated that the PT tested a broader range of skills than other portions of the bar exam tested. Moreover, these newly measured skills are important, indeed essential, to the practice of law.⁷⁰ We have more reason to believe that those who pass a law school test or a bar exam including a PT component are ready to practice than we have for those examinees who are not called upon to demonstrate their understanding of these skills. In addition, it is important not to overlook the learning or educational benefits accruing from the experience of practicing for and taking PT exams. To the extent PT's more closely replicate what lawyers actually do, preparation for PT's should itself enhance the examinees' abilities to practice law. Finally, over time the inclusion of PT's on bar exams may well encourage the teaching and evaluation of lawyering skills in law schools, just as inclusion of PT's on law school exams may well affect the seriousness of law students' study of these skills in their courses.⁷¹

3. *Is the PT a Feasible Alternative?*

As someone who has written traditional law school exams and has experimented in the use of PT concepts in two classes (a first year lawyering process course and an upper class simulation course in negotiating, counseling and interviewing), I can attest to the additional work facing the academic who ventures into this area. PT exam writing is a much more challenging and time consuming task than drafting a traditional essay exam. It involves the compilation of a file consist-

⁷⁰ See text at note 61, and note 61, *supra*.

⁷¹ See text at notes 6-7 *supra*; MACCRATE REPORT, *supra* note 17, at 278-79; NEW YORK BAR EXAM EVALUATION, *supra* note 39, at 3-11 (commenting that "[b]ar examinations are likely to have an effect on law school curriculum").

ing of documents that look like those a lawyer would encounter in working on a matter, such as correspondence, contracts, and investigative memos. The court opinions, statutes, regulations and similar legal materials assembled for the file's "law library" must be collected carefully to ensure that they facilitate a proper solution but without being too directive. And the PT must ring true as a real lawyering task.

If PT's are as valuable as I have argued, then we need to consider in more detail the extent of the time and resource problems they generate, and the ways that these problems might be efficiently solved. So far as law school exams are concerned, these problems do not seem too difficult. Teachers disposed to venture into skills education in the first place may well be willing to put in the time needed to design skills exams too. In addition, the fact that PT's require more work might encourage teachers to work more cooperatively to produce exam questions. (This departure from the usually autonomous and solitary exercise of exam-writing may have other benefits as well.) Furthermore, it may be relatively easy for both law professors and bar examiners to pool their PT's for mutual use and re-use. The California PT's, for example, are made available, together with model answers, as soon as they are administered. Law teachers can use them or modify them as needed. Similarly, in one course I taught, I worked with several other teachers to write a PT. After our first use of the PT, we were able to add or revise fact and law documents relatively easily so that re-use of the core PT was not difficult.

The practicalities bearing on the possible incorporation of PT's on bar exams can also be dealt with. As Dean Harbaugh acknowledges, the costs of PT's are in fact greater than those of other exam methods.⁷² Moreover, simply adding the PT to existing bar exams will by definition lengthen them, and so add to the financial or emotional costs borne by the bar applicant. But it is not actually necessary to add to the total length of bar exams in order to use PT's, nor is it desirable. Rather, the PT could displace some portion of the existing exams, whether in the short answer, essay or multiple choice sections of the exam. As for the administrative costs, these can be addressed, as Dean Harbaugh strongly recommends,⁷³ by having the NCBE assume the responsibility for writing the PT's just as it has for the Multistate Bar Examination (MBE), the Multistate Professional Responsibility Examination (MPRE) and the Multistate Essay Examination (MEE). The result should be to make these costs manageable, and perhaps to reduce them through economies of scale. The NCBE has moved slowly and cautiously regarding the PT, perhaps too cautiously, but it

⁷² See Harbaugh, *supra* note 15, at 13.

⁷³ *Id.* at 14.

has recently, and after extensive study,⁷⁴ made the formal decision to do exactly this.

Beginning in February 1997, the NCBE will offer a model PT (the Multistate Performance Test or MPT), just as it now offers the MBE, the MPRE and the MEE.⁷⁵ These latter tests are now used respectively in 46,⁷⁶ 45⁷⁷ and (as of 1992) 13⁷⁸ states. Whether the MPT is adopted in a comparable fashion remains to be seen. And whether the scope of the skills it tests is expanded to include various other applied lawyering skills also awaits future decisions. But I believe the establishment of the MPT is an important step, reflecting an appreciation that lawyers must possess more knowledge and skills than are now tested on bar exams. Jane Peterson Smith of the NCBE, probably the person most responsible for the development of the PT, suggests that the expanded use of PT's would reflect such factors as the desire of bar applicants to feel that the test is authentic, the public's concern that licensed lawyers be competent and judges' criticism of lawyers' performance.⁷⁹ I would add that the inclusion of lawyering skills on the bar exam will encourage clinical and skills education in law school.⁸⁰

B. *The Video PT*

1. *What Is It and What Does It Test That Is Different from the PT?*

The video PT adds a visual component to the PT in the form of a tape of one or more lawyers performing a lawyering skill in the case for which the PT written materials were provided. The test taker must synthesize the legal and factual materials and integrate them into an analysis of how the lawyer(s)—assumed to be working with the same case file—applied that knowledge in the specific situation presented. That context might be counseling a client or negotiating with an adversary or conducting a cross-examination of a witness.

⁷⁴ In addition to the 1993 study discussed in the text, the NCBE also considered a prototype PT in 1990-91. See AMERICAN COLLEGE TESTING, TECHNICAL EVALUATION OF A PROTOTYPE PERFORMANCE TEST (1991).

⁷⁵ See Robert L. Potts, *Letter From the Chair*, BAR EXAMINER, May 1995, at 2; Smith, *supra* note 52, at 41.

⁷⁶ 1994 Statistics, BAR EXAMINER, May 1995, at 7, 18. The MBE is also administered in the District of Columbia, Guam, Saipan, and the Virgin Islands.

⁷⁷ *Id.* at 21. Like the MBE, the MPRE is administered in the District of Columbia, Guam, Saipan and the Virgin Islands as well.

⁷⁸ Jane Peterson Smith, *The New Multistate Essay Examination*, BAR EXAMINER, NOV. 1992, at 13. (This figure was the number of jurisdictions which, "at this writing," were going to "be administering the July, 1993 MEE.")

⁷⁹ Smith, *supra* note 52, at 42.

⁸⁰ *But cf. id.* (suggesting that increased use of PT's on the bar exam could also reflect law school concerns that they not be pushed too far into mandatory clinical education).

The addition of the visual layer significantly expands the potential of the PT. Clearly it permits the design of tests that measure the subtleties of interpersonal activities more fully than a purely written PT, for the video presents those nuances to the test taker for analysis. In addition, it enables—obligates—the examinee to assess how the performing attorney actually puts all of the pieces together. No critique of a written record of an attorney's work could equally call for an understanding of all of the constraints that affect a lawyering performance, ranging from the client's emotional or psychological disposition, to the time pressures under which the lawyer must operate, to the lawyer's respect for the client's right to make the ultimate decisions.

In this sense, the video PT brings the examinee much closer to engaging in a direct performance of an interpersonal lawyering task than the written PT can. While the test taker does not perform the lawyering role shown on the screen, he or she also does not merely outline a course of action for a lawyer to take to "solve" the PT task, or even prepare a document needed to carry out the task (although the inclusion of such document drafting assignments makes the video PT question an even better reflection of real lawyering). Rather, the examinee is asked to look at the entire picture brought to life by the video, to analyze what worked and what didn't and why, and propose alternative course(s) of action that might have produced better results in the complex, concrete situation presented by the exam. Like the lawyer whose performance is captured on the tape, the student must demonstrate a synthesis of legal understanding and practice skills.

2. *Is the Video PT a Valid and Reliable Test?*

The technique of showing a student a lawyering vignette of one sort or another and then discussing the activity on the tape is being used for teaching in a number of skills-related courses such as interviewing, counseling and negotiating, or trial advocacy. But its use as a testing device seems to be minimal; very few teachers have used video PT's and I am not aware of any bar examination that has included such questions.⁸¹ Thus, it is not surprising that no studies, to my

⁸¹ One of the reasons for the slow pace of development in this area is that while the number of teaching videotapes produced is increasing, production has not yet focused on the particular needs for tapes that can be used for testing. (For a sense of the current state of the field, readers should consult issues of the *Video Law Review*, published by the Media Library at the Social Law Library, Boston, MA.) Even among those involved with continuing legal education, there has only been sporadic interest in developing videotapes either for learning or any other purpose. See, e.g., Larry Smith, *Trends in Video Training: Vendors Gradually Tapping Resistant Market*, 11 PRENTICE-HALL LAW, HIRING & TRAINING REP. No. 1 (1991) (Westlaw Lawprac Index).

knowledge, have been done of the video PT, and that there is no statistical data that might validate the desirability of using this type of exam, either in law school or on the bar exam.

There is, however, a fair amount of anecdotal data that I have collected based on my use of this kind of an exam in two different contexts. Three years of personal experience using such exams gives me some basis for optimism. My assessment is based in part on student response and the views of my colleagues with whom I have taught these courses, and in part on my own subjective observations.

We have used video vignettes of lawyering activities as part of what I would describe as modified PT final exams in two courses, a first year lawyering process course (Lawyering) and an upper class course on negotiating, counseling and interviewing (NCI).⁸² The Lawyering exam is a take-home paper, while the NCI test is an in-class closed book exam.⁸³ In both classes, the students are asked, in part, to write an essay that analyzes the manner in which the lawyer per-

This type of tape production should be encouraged. One of the most important recommendations of the MACCRATE REPORT, *supra* note 17, is to establish an institute for the practice of law. One of the goals of this institute would be to examine the process by which lawyers acquire and refine skills. *Id.* at 33. Development of videotapes and computer materials would clearly fall within its mandate. Unfortunately, the institute has yet to be established.

As I explain in the text, I have been using video PT exams, and so I have become involved in producing the necessary videos myself. I mentioned earlier the additional time it takes to prepare a PT. I should also add the obvious that it takes time to produce the video vignettes discussed in the text. It also requires resources. At our law school, we have an excellent A/V person and some limited editing facilities to assist us in producing tapes. We also have had extensive experience in making tapes of one sort or another in our clinical courses. All of these are factors relevant to the use of such video exams; so is my sense that my colleagues and I have enjoyed making these tapes. I have also used portions of others' tapes (even professionally produced ones) for exam purposes. In every instance, the producer of the tapes readily agreed to their use for this purpose.

⁸² The Lawyering course is a two-credit, first semester required course that *introduces* students to the importance of facts (how to gather them and then distinguish the relevant from the irrelevant) and to some of the other tasks that lawyers perform (interviewing, counseling and negotiating). It also has an intensive introduction to case analysis and stare decisis. The heart of the course involves the students' immersion into an ongoing lawsuit as they work through a 200-page litigation file and a series of video vignettes depicting lawyering activity in the case. The Negotiating, Counseling and Interviewing course is a four-credit elective that uses intensive feedback on five videotaped student simulations and classroom discussion of extensive readings to teach the skills covered.

⁸³ I use the term "modified PT" to describe these exams because in each of these classes I have slightly altered the PT model that is currently being used on bar exams. In Lawyering, the "Law" and "Fact" files for the final exam include extensive documents that the students have received and studied earlier in the semester. In NCI, where the focus is on skills rather than doctrinal learning (in contrast to Lawyering, where both are important), the law component of the final exam is provided in more of a summary fashion than is the case on the bar exam PT, for which the texts of statutes or judicial opinions would be set out. In short, the PT concept is a flexible one and can be adapted to fit the needs of a variety of courses and exam objectives.

formed a particular task. In effect, we ask them to give a written critique of the lawyer's performance. Their analyses should reflect both their various skills readings and our classroom discussions of similar lawyering vignettes.

An important educational point is that the classroom use of video vignettes very much complements the use of vignettes on the final exams. This is particularly so in the Lawyering class, where the students first encounter the lawyer seen on the final exam tape in another video seen in class. On the exam tape, that same lawyer is performing different tasks in the same case with which the students already have acquired detailed familiarity. Conversely, the students' advance notice of the exam mode increased their attentiveness and immersion in the classroom analyses of the tapes. As I suggested earlier in this Article, what happens on the final exam affects what happens (and works) in the classroom—and vice-versa.⁸⁴

While our experience is not a scientific study, I think it is fair to say that the students' written critiques of taped lawyering vignettes do reflect in part their skills abilities. The final exams certainly do not comprehensively assess students' skills, but they do provide some indication of students' awareness of what is entailed in a lawyer's interview or counseling of a client. The videos seem real to the students. The vignettes thus seem effective in triggering thoughtful analyses of what the lawyer is trying to do.⁸⁵

Moreover, in both courses we have sought to enhance the tests'

⁸⁴ See text at notes 6-7 *supra*.

⁸⁵ Video vignettes also can be quite useful in testing doctrinal knowledge. For example, in civil procedure, asking a test taker to view deposition vignettes and then answer an exam question on the efficacy of a particular discovery rule or on the utility of discovery sanctions should produce richer doctrinal analyses than a strictly paper exam. The analysis of such a visual depiction can take into account the statutory language and relevant case law underlying the doctrine, as well as the lawyers' expressions, the language of their verbal exchanges and their more subtle communications (to each other and to their clients). Watching a lawyer make a "speaking objection" to a deposition question, for example, can vividly illuminate the rationale for sanctions against discovery abuse. (A "speaking objection" is a long-winded objection made by a deponent's lawyer for the purpose of advising the deponent how to respond to the question.)

Similarly, in a mental health law exam, a video vignette of a lawyer-client counseling session could be the basis of a question on the substantive law applicable to competency determinations. Professor Michael Perlin of New York Law School frequently uses such questions as part of his exams in his mental health law courses. Students could also be asked to evaluate the lawyer's counseling skill on the basis of the same videotape. Evidence is another course that video vignettes could illuminate from a student's learning perspective; the bases of sustaining objections are much clearer when observed in context. The same is true for any exploration of ethics dilemmas, as illustrated in the valuable videotape materials of STEPHEN GILLERS, *ADVENTURES IN LEGAL ETHICS—A VIDEOTAPE PRODUCTION FROM NEW YORK UNIVERSITY LAW SCHOOL* (1993) (and the accompanying *Teachers' Guide*).

validity by insuring that students understand and address the standards on which they are being evaluated. To that end, we provide the students in advance with criteria meant to provide a framework for the student critiques. Just as the clinician's critique is based on generally accepted standards of skills performances, the same standards are (or at least are supposed to be) the bases of the student critiques, and then can be the basis of the teacher's evaluation of the students' conclusions.⁸⁶

In particular, the good answer analyzes why something the lawyer did either worked or did not work. Of course, everything is relative. The level of skills analysis of first year students is usually elementary at best. The critiques in the NCI final exams, on the other hand, are often quite good. In this four-credit course, the final exam critique provides the students with an opportunity to bring to bear a semester's study of interpersonal lawyering skills. Their analyses often reflect that coherent whole. Just as the lawyer being watched must synthesize law and execution, so the student evaluation must reveal a parallel understanding of that integration process.

In both the first year and upper class courses the video PT is only a portion of the final grade.⁸⁷ My unscientific correlation of the students' NCI exam grades and their grades on their simulation exercises (based on teacher observation) indicates that generally the students who do better on the written critiques also do better in their performances. This is the kind of study that our psychometrician colleagues would say has not yet been completed in any thorough way. I agree. But the need for more study should not deter us from continuing to try out these new techniques; on the contrary, we need to try these techniques in order to assess them.⁸⁸

⁸⁶ There has, however, been some student resistance to these written skills critiques. There are those among the student body, faculty and the bar generally who feel there is simply too much subjectivity inherent in the process of evaluating such critiques. For reasons explained throughout this Article, I respectfully disagree. As long as there is appropriate advance notice of what the criteria are, *see* note 25 *supra*, and as long as the bases for those criteria have been studied in the course, students should be able to write, and faculty to grade, as rigorously as would be the case on any essay exam.

⁸⁷ In Lawyering the lawyering skills critique constitutes 25% of the exam grade; the remainder is based on the students' analyses of what the lawyer should have concluded about the merits of the case in the updated factual and legal circumstances laid out in the exam question. In NCI, the final exam constitutes 35% of the grade; the remainder is based on graded simulation exercises and classroom discussion.

⁸⁸ As with any exam, we should try to be as fair as possible in our design and grading. Where the consequences of any unfairness are great, as in the bar exam or other "high stakes" exams, it makes sense to innovate cautiously, but we should not be deterred from trying out new and promising testing methods, especially in individual law school courses, because definitive psychometric evidence has not yet been gathered.

C. *The Interactive Video Exam*

By integrating the video image with the capacity of the typical personal computer, we now have the technology to enable a user to interact with the events shown on the screen.⁸⁹ This interactive function offers new ways in which persons may learn and also new possibilities for evaluating what a student knows. While the technology has been around for some time,⁹⁰ and is being exploited actively for training and educational purposes in a number of areas, it has had only a limited impact thus far on law and legal education.⁹¹ In a

⁸⁹ We can also enable computer users to interact with written material presented to them through the computer. Non-video computer law study programs have been developed and are in use in a number of subjects. Among the leading developers of such programs has been the Computer Assisted Legal Instruction (CALI) program at the University of Minnesota.

Some software now permits the user to explore questions further in "hypertext," through which the user can delve more deeply into the issues surrounding a particular concept shown on the screen by easily opening sub-libraries contained in the software. "Hypertext"—a term coined to mean nonsequential writing, Evelyn Richards, *Computer Technology for the Wandering Mind*, WASH. POST, May 1, 1989, at F1—is, in today's computer applications, a way of allowing users to skip freely from a key word or phrase to much more information on that particular subject. Richard O'Reilly, *An Improved Program for Making Free-Form Database Files*, WASH. POST, Sept. 13, 1993, at F21. Hypertext allows the user to jump from one information source to another without having to use keyboard searches. Jonathan Weber, *Autodesk Faces Dicey Task of Broadening Line Software*, L.A. TIMES, Dec. 30, 1990, at D1.

⁹⁰ In fact, the basic technology has been developed and available to the public for a decade. See Ellen Miller, *Fusion of Computer and Video Creates Novel Learning Tool; Develops Practical Legal Skills*, NAT'L L.J., SPECIAL SECTION, CONTINUING LEGAL EDUCATION, Mar. 24, 1986, at 15. It seems likely that the rapidly developing CD-ROM technology will make it much easier to use these techniques than has been the case with programs using video laser disks, which require much more cumbersome hardware than do CD-ROM's.

⁹¹ For a discussion of the use of multimedia approaches and interactive video in such diverse areas as medicine, adult literacy, history, and archaeology, as well as law, see Vicki L. Reeve & Alex S. Kassten, *SALT [Society for Applied Learning Technology] 92-14th Annual Conference: Interactive Systems for Training, Education, and Job Performance Improvement*, 10 MULTIMEDIA & VIDEODISC MONITOR No. 10 (Oct. 1, 1992) (Westlaw MVIDEOMON database). Although these programs have begun to be used in law schools, or at least have been made available to students and teachers, their primary users so far appear to be people responsible for continuing legal education. See generally American Institute for Law Training Within The Office (AILTO) & Association of Legal Administrators, *In-House Training: Maximizing Your Lawyers' Professional Potential—ALI-ABA Course of Study*, Feb. 18, 1994, at 49, 53, 86 (unpublished materials, on file with the author); Stephen T. Maher, *Interactive Video Continuing Legal Education: Future and Present*, CLE J., Nov. 1995, at 10.

In an interesting example of creative lawyering and of the use of this technology in a different area of legal education, attorneys at Skadden Arps, Meagher, Slate and Flom produced an interactive program for the employees of their corporate clients to test the employees on what constitutes inappropriate behavior for purposes of preventing sexual harassment claims. Saundra Torry, *Computers, Theatrics Dramatize Harassment Issues*, WASH. POST, June 20, 1994, at F7.

number of fields, however, and in some aspects of legal training, interactive video is being used for teaching purposes, and this section begins by describing the ways that interactive video can contribute to student learning. Then I will turn to the more speculative, but promising, question of how this technology could be employed in evaluations as well.

What happens in interactive video learning? An individual user sits at a computer monitor, observes various scenes shown on it and then responds to questions about what is being viewed.⁹² Depending on the user's responses and the software's sophistication, the exercise might lead the user in a number of different directions. Suppose, for example, that the user is viewing a trial on the video screen—a likely example, since the two areas that have seen the most development in law have been programs in evidence and trial advocacy.⁹³ As testimony is being taken, the user is asked to make appropriate objections. The "court"—that is, the program—then rules on the objections and, depending on the responses, the program either continues with the trial, or pauses to explain to the user why his or her objection was unfounded. There is considerable evidence that such programs do help people to learn.⁹⁴

⁹² Harvey Dale, *Innovation: Computers and Computer Technology, in CLE AND THE LAWYER'S RESPONSIBILITIES IN AN EVOLVING PROFESSION, THE REPORT ON THE ARDEN HOUSE III CONFERENCE, NOVEMBER 13-16, 1987, at 277 (ALI-ABA Committee on Continuing Professional Education, 1988)*. Dale's excellent overview of the technology and its potential in legal education, though written several years ago, is still a useful and timely summary.

⁹³ Thus, for example, seven out of eight video "learning modules" in a laser disc program developed at Stanford teach evidence or trial skills, while one teaches client interviewing skills. See *THE INTERACTIVE COURTROOM, LEGAL SKILLS TRAINING* (Stanford Law School Interactive Video Project 1990). Similarly, 15 of 21 interactive videos developed at Harvard teach evidence or trial skills; others focus on negotiation techniques, client interviewing skills, ethics, and the dynamics between lawyer, client, and opposing party. *INTERACTIVE VIDEO LIBRARY* (Harvard Law School 1992).

⁹⁴ Some studies have shown that interactive computer users learn more than those who do not have access to the technology. See, e.g., Delores Kong, *Video "Patients" Hone Student Skills in Real-Life Way*, *BOSTON GLOBE*, Mar. 11, 1991, at 37 ("Third year medical students who used video technology scored an average of 7% higher on exams than those who did not . . ."). In a much different area—hotel management—the use of interactive video as a learning device produced a 25% improvement in retention. *New Training Looms; Interactive Education*, *HOTEL & MGMT.*, Apr. 4, 1994, at 26.

One reason that these programs can be effective is that they can take advantage of the privacy of the locale of the user's computer/video terminal. The user can take the lessons as speedily or as slowly as the user wishes. Provided the program is constructed to ensure privacy, no one need know either how long it took the user to complete the program, or how well the user performed. This element of privacy can reassure and assist the program's users. Thus it may be that, as a continuing legal education professional in Canada recently observed, interactive video is "best used as a stand-alone product," one that does not require either an instructor or a class setting. David Cruickshank, *Supplemental Materials on Implementing a System for Lawyer Training At Your Office*, ALI-ABA

As in other aspects of experiential learning, medical educators are far ahead of their law school counterparts in the use of interactive video.⁹⁵ Such programs are in use in the training of physicians, nurses and other medical personnel. Medical educators have concluded that the interactive program “provides an introduction to what students can expect as part of their clinical experience—it is as close to real life as possible.”⁹⁶ Instructional modules cover such diverse areas as reading blood slides or linking the visual appearance of human movement disorders with particular scanning images of brain disorders. Especially relevant to legal educators are the interactive programs involving taped patient interviews in which the key issues are whether appropriate information was obtained for purposes of ordering various diagnostic tests.⁹⁷ As in the legal education evidence example given above, the medical students are given the opportunity at regular intervals to compare their decisions to those of an expert, as reflected in the program.

While a principal purpose of interactive video programs is to facilitate the users' self-learning, they clearly also have significant potential as testing devices. Conceptually, the computerized feedback on self-learning exercises is the same kind of evaluation that grading and critique of an exam answer would constitute. In the context of skills education, moreover, the technology provides an experiential form of learning that brings the user as close as possible to the actual doing of the task in question, short of directly enlisting the student in either a simulated or real performance. Certainly there are questions of validity and reliability to be answered for these exams, as for all exams, but in principle their validity should be quite comparable to that of the video PT. Moreover, the use of the computer may insure that the interactive exams' reliability or consistency is greater than that of exams graded by humans (though this consistency may come at a price in terms of rigidity of the program's specifications of what answers are “right” and “wrong”).

I do not suggest that “the traditional teaching style of formal lectures and tests will die off . . . and be replaced [through computerization] by [a] more interactive learning style,” as one leading educator

Course of Study, Feb. 18, 1994, at 126 (unpublished materials on file with the author). As I will argue, however, these programs also have promise in the much less private sphere of testing.

⁹⁵ See, e.g., *An Electronic Curriculum? Multimedia Health Education*, WORLD HEALTH, Jan. 1994, at 26; Thomas M. McMillan, *Interactive Video Bolsters Health Care Education: Includes Related Article on Multimedia Products From Different Vendors*, COMPUTER GRAPHICS WORLD, May 1990, at 44.

⁹⁶ *Id.*

⁹⁷ *Id.*

has predicted.⁹⁸ The objective here, as in the use of interactive video programs for learning purposes generally, should not be to supplant traditional testing methods, but to supplement them, and to maximize the efficient use of technology in ways that enhance the ability to evaluate someone's competence. Unfortunately, as I have already noted, the world of legal education lags behind in its exploitation of available technology. Even the quite modest step of placing conventional standardized tests on line has not progressed far.⁹⁹ While interactive video testing is still a developing field, there would seem to be no reason why this technology could not be employed to assess a test taker's competence. Such work is in progress in other professions,¹⁰⁰ and there are attractive possibilities for its use in the law.

Computerized video testing seems especially appropriate for the various lawyering skills that are not typically tested on bar exams and in law school final exams, such as counseling or interviewing or negotiating. There would seem to be two basic ways in which the interactive programs might work. One is much more complicated than the other to design, and it may be more than is realistically feasible, especially in the short run. It would ask the user to make choices at various points during the viewing of a lawyering activity and then (via the computer) to direct the person on the videotape to execute those

⁹⁸ This is the view of James B. Appleberry, President of the American Association of State Colleges and Universities, as paraphrased in Larry Gordon, *Imagine the Class of 2013*, L.A. TIMES, Feb. 9, 1993, Special Section at 3.

⁹⁹ Outside the field of legal education, the Graduate Record Examination (GRE) will soon be given entirely on computers. The test takers will not only dispense with the number 2 pencil, but will be able to obtain their scores immediately after taking the exam. Michael Winerip, *No. 2 Pencil Fades as Graduate Exam Moves to Computer*, N.Y. TIMES, Nov. 15, 1993, at A1. Although students will still have the option of taking the GRE by paper and pencil, the normal route, by 1996-97, will be on computer.

I am not aware of any movement to put the LSAT on line. Nor am I aware of any effort to computerize either the various parts of the uniform components of the bar exams (such as the MBE, MEE or MPRE) or any of the local state portions of the bar exams. While computerization would seem to be relatively easy for the short answer exams, it also seems feasible for essay questions to be administered and answered on computer, provided that the test taker is able to write on a computer keyboard.

It is worth noting, however, in light of the substantial criticism of standardized multiple choice exams, that "[s]imply automating a bad test does nothing to solve the problems of a bad test," *Id.* (quoting Cynthia Schuman, Director, National Center for Fair and Open Testing). Moreover, computerized testing (like computerization generally) has its limitations, including, as might be expected, security and technical problems. See William Celis III, *Computer Admissions Test Found to Be Ripe for Abuse*, N.Y. TIMES, Dec. 16, 1994, at 1; William Honon, *Computer Admission Tests to Be Given Less Often*, N.Y. TIMES, Jan. 4, 1995, at A17.

¹⁰⁰ See, e.g., Kurt Pitzer, *Pierce College Gets Grant from Kaiser*, L.A. TIMES, Apr. 20, 1994, at B3 ("nursing students will use interactive computer programs to test their skills"); *Proposed Changes in the National Teachers Examinations*, PHILA. INQUIRER, Nov. 13, 1988, at H20.

choices.¹⁰¹ Such a program would require the test-taker “virtually” to perform, and would also press the examinee to address the impact of context as he or she coped with the results, at later stages of the video, of the choices made earlier on.

A second, more easily implemented computer video exam would simply ask the user to answer various questions about what is viewed on the video screen. The video vignette would stop at the point for a question and continue after the answer is given. The actual questions could be multiple choice, short essay or full-length essay questions, and the answers could be inputted on the computer in each case. This type of exam is nearly identical to the concept of video PT exams discussed above. Like them, it would directly test legal knowledge¹⁰² and/or critique skills, but not actual performing skill. Also like the video PT, however, this interactive PT would bring the test-taker close to the realm of performance. The difference between this test and the ordinary video PT would lie in the examiner’s ability to focus the test-taker’s attention on particular moments, and particular issues, in the video vignette.

In fact, such programs already exist in the legal education area, although not in great numbers. For example, one of the evidence programs has a feature that prints out the user’s performance results, indicating the time spent on the lesson and the percentage of issues spotted and correct responses given.¹⁰³ As with interactive video technology generally, the testing capacity has been available for a number of years.¹⁰⁴ For reasons that are not entirely clear, there has been little progress in legal education toward implementing these concepts. I believe it is time for us to correct this omission.

¹⁰¹ Such branching, based on user decisions, would require the computer program to offer different tracks on which the lawyering interactions might proceed, each of them reflected on video which the program could display. Branching is already available in some of the interactive trial advocacy/evidence video programs. In other skills contexts where there may be more than one accepted course of action—mediation, for example—the challenge is to design a program that creates a “variety of possible scenarios, using limited branching.” Reeve & Kasten, *supra* note 91 (paraphrasing Jeremy Seeger of Harvard Law School).

¹⁰² One evidence teacher uses commercially produced interactive video programs to test his students in an advanced evidence seminar on doctrinal issues. Interview with Professor Robert Pitler, Brooklyn Law School, July 12, 1994.

¹⁰³ Kathy M. Morris, *Interactive Video: One Solution to the Multi-Office Training Challenge*, 13 PRENTICE-HALL LAW. HIRING & TRAINING REP. NO. 4 (1993) (Westlaw Lawprac Index).

¹⁰⁴ The technology has in fact been available for some 10 years. Tim Hallahan, a leading developer of these programs, commented a decade ago that “[t]he computer can be a testing, as well as a learning, vehicle. Because the lessons are structured, it might be important for the firm to track who took the lessons, completed them and how well they did.” Miller, *supra* note 90, at 15 (quoting Hallahan).

CONCLUSION

As teachers of lawyering skills, we have an obligation to look at how we determine whether our students have absorbed our teaching. For those of us who teach in live-client clinics, our intuition always has been that direct observation and feedback are the best, if not the only, way to evaluate our students. I think that instinct has for the most part been correct. Even in the clinic, however, there may be ways to supplement our observations so as to enhance the fairness of our evaluations, without compromising the personalized nature of clinical teaching.

But increasingly lawyering skills are being taught in non-clinic contexts. In those situations, little attention has been given to whether or how students have incorporated that learning. Both in non-clinic law school classes, and, importantly, on bar examinations, we must address the question of whether to evaluate students (or bar applicants) with respect to those skills. For me, the answer is clear: we must evaluate students' skills or risk devaluing the skills being taught and (especially in the context of the bar exam) perpetuating the myth that passing exams on legal doctrine demonstrates competency to practice law.

How then should we do these evaluations? Observation and direct feedback remain, in general, the best way to evaluate students' oral lawyering skills. Many of us and our non-clinical colleagues, however, teach these skills in varying ways and in varying degrees in a non-clinic context, where direct observation by full-time faculty is simply not possible. Even observation by others such as adjuncts (whatever the virtues and problems of this system) is certainly not available on bar exams. While written tests are not perfect substitutes for clinical evaluation, the basic concept of the performance test, as well as its further development in the video PT and the intriguing possibilities for the interactive video PT, all suggest that the need for new evaluation systems can be met. There is good reason to believe that some or all of these methods are actually valid and reliable techniques. We should use and experiment with these evaluation methods and continue the efforts to refine them.