Using Examinations in First-Year Legal Research, Writing, and Reasoning Courses*

Douglas Miller**

I. INTRODUCTION

Examinations are typically not used in first-year legal research, writing, and reasoning (LRWR) courses; there are some legal writing programs that use them, but these programs appear to be in the minority.[1] I thought it might be useful to discuss some of the reasons for this, to identify some of the objections typically put forward to the use of these examinations, and to see whether I could develop some responses to the objections. The discussion proceeds by means of three straightforward propositions, but it also includes references to actual exam questions that have been used either in the legal writing program at South Texas College of Law, where I teach, or elsewhere. (Ca-

---

* This article, like the others in this issue of *Legal Writing*, is based on a presentation given at the 1996 Legal Writing conference sponsored by the Legal Writing Institute. This presentation took place in Seattle on July 20, 1996. In preparing this article, I have retained the tone of the original presentation, which was in essence an exhortation for those teaching legal writing, even though I recognize that not all readers of this article will be legal writing teachers. I wish to thank the Legal Writing Institute for providing a forum for this presentation, and for putting out this special *Proceedings* issue. Finally, I wish to thank Diane Edelman, Pamela Lysaght, and David Walter for their thoughtful editorial suggestions.

** Assistant Professor of Legal Writing, South Texas College of Law, Houston, Texas; J.D., Boalt Hall School of Law, University of California at Berkeley, 1983; B.A., University of California at Santa Barbara, 1978; admitted to practice in Pennsylvania and Alaska. (My colleague Kim Cauthorn of South Texas College of Law was originally scheduled to participate in the LWI presentation in July, but unfortunately was unable to do so for medical reasons.)

[1] A decidedly unscientific survey has turned up twelve law schools that use something that could be described as an examination in at least one LRWR course. Those schools are the University of Detroit Mercy School of Law, Mercer University Law School, Case Western Reserve University Law School, University of South Carolina School of Law, Stetson University College of Law, South Texas College of Law, Chicago-Kent College of Law, University of Mississippi School of Law, Loyola University School of Law (New Orleans), Nova Southeastern University Shepard Broad Law Center, Saint Louis University School of Law, and Albany Law School. At two other schools (McGeorge School of Law and the University of Cincinnati College of Law), apparently at least one instructor currently uses an examination but others do not. See note 4, *infra*, for more detail about the examinations used in some of these programs.

veat: some of the questions are problematic, and have been included only for purposes of discussion or illustration.)

II. EXAMINATIONS HAVE VALUE

Most law school courses use final examinations. Most undergraduate courses in this country still use examinations. Examinations provide a means of ensuring that students who receive academic credit for a course have mastered the content of the course, and this holds true regardless of whether the content of the course is best described as a set of skills, a set of principles, or a body of knowledge. Furthermore, giving an examination provides an instructor with a means of distinguishing among students who have attained a bare mastery of the contents of the course, students who have attained a complete mastery of the contents of the course, and students who fall somewhere in between these levels of mastery. Finally, giving an examination increases learning in two distinct ways. First, the process of preparing for an examination increases learning: instructors who use examinations can attest that on the day of the examination the vast majority of students understand more about the content of the course than these students did at the end of the last scheduled class, as a result of the preparation process. Second, the act of taking the examination itself provides an educational experience, as illustrated by the fact that most students know more at some point *after* the examination (whether *immediately* after the examination, after a few days *ruminating* about the examination, or after participating in an organized *review* of the examination) than they did immediately before the exam.[2]

---

[2] Perhaps a better way to say this is that students will understand more after the examination than they would have had they never taken the examination. We have all had the sensation of information or knowledge draining steadily out of our brains in the hours, days, or weeks following an examination. The common wisdom arising from this sensation is that we actually know less after the examination than we knew before it, which might require the modification of phrasing I have made here. But this common wisdom may itself be open to challenge: perhaps the knowledge or information merely has been transferred to a less accessible part of the brain, and can be recalled under the proper stimulus or conditions. For example, the designers of review courses for bar exams assume that there is a well of information that an instructor can draw upon, at least in subjects that are required in law schools, even though the common wisdom might be that students have forgotten much of what they once knew. Or perhaps we can say that the goal of our courses is actually to have students develop certain skills, and these skills have been enhanced by the experience of the examination, to be retained by the student over the long term, unless allowed to deteriorate from lack of use.

The second use of examinations mentioned above, distinguishing among students, is not applicable to ungraded or pass/fail courses. Yet the vast majority of LRWR courses offered in this country, like the vast majority of non-LRWR courses, are not offered on this basis,[3] so we can safely retain this part of our description of the value of examinations.

III. MASTERY OF THE CONTENT OF OUR COURSES CAN BE TESTED UNDER EXAMINATION CONDITIONS

This proposition is more controversial. Most law school courses have examinations of a duration that corresponds to the number of credit-hours given, so that a typical three-credit-hour course would have a three-hour exam. Many education theorists would say that some skills cannot be adequately tested in three hours, and some LRWR teachers are quick to point out that some of the skills we teach fall into this category. For example, it would be difficult to argue that the skill of investigating a set of complex legal issues in a law library can be tested in a three-hour, in-class examination.

Yet there are several answers to this observation. One is that this point hardly renders examinations themselves invalid: for every skill that *cannot* be tested in three hours, there are many that *can* be tested in three hours. The education theorists who conclude that not every skill can be tested in three hours do not advocate the abolition of examinations: rather, they suggest that in-class examinations be used with other kinds of projects. Thus, if we conclude that there are skills taught in our courses that cannot be tested in three hours, it does not necessarily follow that examinations are useless for us, for the simple reason that we might use in-class examinations in conjunction with other kinds of assignments.[4]

---

[3] Responses to a 1994 survey suggest that 75% of legal writing programs use letters or numbers to grade student performance, and another 8% of programs use an "honors/pass/fail" system; only 11% of programs appear to use a pure pass/fail system. Jill J. Ramsfield, *Legal Writing in the Twenty-First Century: A Sharper Image*, 2 LEGAL WRITING 1, 75 (1996).

[4] For example, in the first-semester LRWR course at South Texas, students are required to complete many non-examination assignments, including an open research memorandum of law, and the in-class examination usually counts for only about 40% of the grade. At Albany the examination has roughly the same weight. At Mercer an examination counts for approximately 25-35% of the grade in the second-semester LRWR course. At Saint Louis a research examination given early in the second semester counts for 10% of the five-credit, year-long LRWR course. At Loyola (New Orleans) an examination on research and citation counts for 25% of the first semester LRWR grade. At Case

Another answer is that not all examinations need to be of limited duration, nor do they all need to be administered in a classroom. So-called "take-home" examinations have been used for decades in doctrinal courses at law schools around the country. Some of these have very strict requirements regarding duration and sources to be consulted. For example, a student might be required to pick up the exam at 9:00 a.m., and turn in an answer at 5:00 p.m. that same day, or perhaps 9:00 the next morning. This arrangement allows the instructor to measure a broader range of competencies, without some of the disadvantages associated with more traditional kinds of "papers." Some of the skills we teach could be adequately tested by giving a more traditional memorandum of law assignment, but limiting the duration and sources to be consulted or tailoring the end-product requirements. (Some readers might object that these could not, strictly speaking, be called "examinations." It hardly matters whether we call these assignments "memoranda" or "examinations," so long as we agree that they are not currently used in many programs, and that they might prove useful.) Thus, even library research skills might be tested by giving students a fact pattern and a time limit, and asking them to come back with a list of the ten or fifteen authorities that appear to be most likely to prove helpful in resolving the legal questions posed by the fact pattern, along with a brief explanation of the significance of each authority, and perhaps a narrative of what steps and missteps were taken along the way.[5] While there might be practical problems to be solved before an instructor would decide to employ this means of evaluation, this does not undercut the main point: that mastery of skills in our courses might be tested by means other than traditional writing assignments.

Once we begin to analyze the content of our courses with the goal of identifying discrete skills, we can begin to imagine

---

Western Reserve an examination on fundamental legal research counts for 10-15% of the grade in a year-long course. At Nova a legal research and citation examination counts for 10% of the grade each semester. At Detroit Mercy, the examination counts for 18% of a five-credit, two-semester course. Chicago-Kent uses an examination on research that counts for 10% of the second-semester LRWR grade. South Carolina uses an examination to determine the entire grade for its one-credit research course. Stetson uses an examination on research and citation and 15% of the grade is based on the examination.

[5] Mary Jensen of the University of Mississippi School of Law apparently uses something similar to this. Mary B. Jensen, *"Breaking the Code" for a Timely Method of Grading Legal Research Essay Exams*, 4 PERSPECTIVES 85 (1996).

ways that these discrete skills might be tested in timed examinations. I submit that the range of skills that *can* be tested in this way is remarkably broad, and the range of skills that *cannot* be tested in this way is quite narrow. (And note that even if this were not true, there might still be a place for exams in LRWR courses, for the reasons explained above.)

A few examples might help here. One skill that we all might agree is important is the ability to recognize relative ranks of authority, or at least to distinguish binding authority from non-binding authority, without knowing anything about the authority other than the information contained in a citation. (This is an important part of research as well as analysis.) This skill in turn may be dependent upon an understanding of the structure of court systems. Thus, on an examination, a student might be asked to do one or all of the following:

> **1. Diagram the different levels of the federal court system and explain the basic functions of the courts at each level.**
>
> **2. Diagram the different levels of a typical state court system and explain the basic function of the courts at each level.**
>
> **3. Diagram the structure of the Oklahoma court system.**

More directly, a student might be asked to distinguish binding from non-binding authority, and to demonstrate an understanding of the bases for making these distinctions, in a multiple-choice format:

> **You are trying a lawsuit in a Texas District Court in El Paso. One of the issues involves a federal statute. The trial judge has asked for trial briefs on how to apply the federal substantive law. Your research for the trial brief has uncovered the following authorities that address your issue. In the following questions, circle the statement that *best* explains the authoritative value of each authority.**
>
> **4. An opinion from the Houston [1st Dist.] Court of Appeals is**
>
> **a. binding, because it is from a higher court.**
> **b. not binding, because it is not from a higher court.**

c. not binding, because El Paso is in a different appellate district.
d. not binding, because the issue is one of federal law.

5. An opinion of the United States Supreme Court in an appeal from a decision of the United States Court of Appeals for the Second Circuit is
a. binding, because it is from a higher court with direct appellate jurisdiction.
b. binding, because the issue is one of federal law.
c. not binding, because El Paso is not located in the Second Circuit.
d. not binding, because a state case would not go to the United States Court of Appeals.

6. An opinion of the United States District Court for the Eastern District of Texas is
a. binding, because the issue is one of federal law.
b. binding, because El Paso is in the Eastern District.
c. not binding, because there is no appellate or revisory jurisdiction.
d. not binding, because El Paso is not in the Eastern District.

7. An opinion of the United States Court of Appeals for the Fifth Circuit, in an appeal from a decision of a United States District Court in Louisiana, is
a. binding, because the issue is one of federal law.
b. binding, because Texas is in the Fifth Circuit.
c. not binding, because Texas is not in the Fifth Circuit.
d. not binding, because there is no appellate or revisory jurisdiction.

8. An article in the Texas Law Review is
a. binding, because this is the most prestigious law review in Texas.
b. binding, because it is published by a governmental entity.

c. not binding, because it is not primary authority.
d. not binding, because students from El Paso are more likely to attend Texas Tech than the University of Texas.

9. An opinion of the Texas Supreme Court is
a. binding, because a case from El Paso could conceivably end up in the Texas Supreme Court.
b. not binding, because the Texas Supreme Court has only discretionary review over some matters.
c. not binding, because the issue is one of federal law.
d. not binding, because the case could only reach the Texas Supreme Court by way of a writ of mandamus.

Note that some of the questions above draw upon particular reading or class discussion. For example, in the early weeks of the first-semester LRWR course, students in my sections read a brief explanation of the doctrine of precedent; I use this explanation, even though it is somewhat idiosyncratic, because I like the way it divides the doctrine into different strands. Question 5 draws on our class discussion of these strands. Others might disagree about the phrasing of that question, but it should make sense to students in my course. Similarly, Question 7 is tied to some very specific reading in the course: because we read decisions in which Texas courts have explicitly held that they are not bound to follow decisions of the Fifth Circuit, even on questions of federal law,[6] students in my course should be able to answer this question. I see nothing illegitimate about drawing on specific reading or class discussion, even though students who did not take my course might not be able to answer these questions, or might answer them differently, so long as I am not setting forth pure claptrap as gospel. Unlike, say, state bar examiners, we are not under any obligation to make certain that our

---

[6] *See* Penrod Drilling Corp. v. Williams, 868 S.W.2d 294, 296 (Tex. 1993); Summertree Venture III v. FSLIC, 742 S.W.2d 446, 451 (Tex. Ct. App. 1987). (At least one current United States Supreme Court Justice believes that this principle applies in all state courts. Lockhart v. Fretwell, 506 U.S. 364, 376 (1993) (Thomas, J., concurring) ("[a]n Arkansas trial court is bound by this Court's . . . interpretation of federal law, but if it follows the Eighth Circuit's interpretation of federal law, it does so only because it chooses to and not because it must").)

questions are so generic as to be answerable by any well-versed student from any law school in the state. Instructors in doctrinal courses routinely take advantage of this opportunity to test students on their understanding of material specific to that course, taught in that semester, by that professor.

Some questions might even refer directly to specific reading or class discussion:[7]

> **11. Professor Neumann divides the potential lifecycle of a civil case into six phases. What are they?**
>
> **12. Karl Llewellyn said that, in a sense, no case can have meaning by itself. What did he mean?**
>
> **13. According to the majority opinion in *Payne v. Tennessee*, all other things being equal, should a court be more willing to overrule its own prior decision on a question of *statutory* interpretation or a question of *constitutional* interpretation? Why?**
>
> **14. How does Harry W. Jones, in his article on the common law, explain the persistence of common-law ways of thought in an age of statutes?**
>
> **15. Explain the basic position of United States Supreme Court Justice Antonin Scalia with regard to the use of legislative history in interpreting statutes.**
>
> **16. As we discussed in class, the law sometimes permits both a civil and a criminal cause of action based upon precisely the same conduct and result. Explain why this is allowed, and mention a few of the ways in which the two proceedings would differ.**
>
> **17. Is legal reasoning more like deductive reasoning or inductive reasoning? Explain your answer.**

Let us consider some other skills that are sometimes assumed to be testable only in the context of a large, long-term

---

[7] Works referred to in questions 11-15 include Richard K. Neumann, LEGAL REASONING AND LEGAL WRITING 8-13 (2d ed. 1994); Karl Llewellyn, THE BRAMBLE BUSH: ON OUR LAW AND ITS STUDY 55-65 (1960); Payne v. Tennessee, 501 U.S. 808, 828 (1991); Harry W. Jones, *Our Uncommon Common Law*, 42 TENN. L. REV. 443, 454-59 (1975); and Elizabeth A. Liess, Comment, *Censoring Legislative History: Justice Scalia on the Use of Legislative History in Statutory Interpretation*, 72 NEB. L. REV. 568 (1993). The classroom discussion relevant to question 17 is suggested by works such as Steven J. Burton, AN INTRODUCTION TO LEGAL REASONING 25-78 (2d ed. 1995), and Charles C. Calleros, LEGAL METHOD AND WRITING 56-60, 120-25 (2d ed. 1994).

writing assignment. Drafting of formal issue statements or questions presented might be put in this category. But a question might be constructed that gets at this skill:

> **18. Brazoria County ordinance number 38 provides that "no illuminated signs shall be employed after 2:00 a.m. in the county."**
>
> **Peter Halsey operates a drugstore in Brazoria County. Halsey keeps two lights mounted on a 20-foot pole in front of the store turned on all night even though the drugstore closes at 7 p.m. The lights are directed toward the parking area, but because of the lights, the word "Halsey's" painted in a reflective substance on the front of the drugstore can be seen from a distance of approximately 400 feet.**
>
> **Authorities repeatedly warned Halsey not to leave the lights on past 2:00 a.m., but Halsey refused to turn off the lights. Eventually, Halsey was cited by the county sheriff for violating the ordinance.**
>
> **Courts have held that the illumination of names or numbers on the front of residences and residential mailboxes does not violate the ordinance. The county admits that no citizen complaints were ever filed about the lights, and that there are no residences within ¼ mile of the drugstore.**
>
> **State the main issue in this case.**

Of course, there are those who may object that, while it may be true that some skills can be broken down and tested independently, the tasks that are the main focus of the course, solving legal problems and expressing the solutions in writing, are somehow more than the sum of their parts. In other words, those voicing this objection would maintain that the ability to join all the discrete skills together, which might *itself* be described as a discrete skill, cannot be tested in a limited-time setting. But I do not concede that this is necessarily true. Certainly, state bar examiners do not concede the truth of this objection. Consider the following paragraphs:

> **I need your help preparing for the trial in this case. As you may know, we represent the plaintiff, Esther Dodson, in this product liability action arising out of an automobile accident in which our client was the driver and her sister, Denise Johnston, was a passenger. Denise claims she was injured in the accident,**

> **and we have reason to believe that Denise intends later to sue Esther for negligence.**
>
> . . . .
>
> **There is evidence that, following the accident, CEC changed the design of the sunroof. We fully expect that CEC will attempt to prevent this information from getting into evidence. There may be a conflict between Rule 407 of the Federal Rules of Evidence and the Columbia Product Liability Act concerning the admissibility of subsequent remedial measures in product liability cases. Please do the following:**
>
> **Write an objective memorandum for me discussing which of the statutes will be applied, the theories supporting admission of this information, and the likelihood that each theory would be successful. In order to assist you, I have attached Rule 407 of the Columbia Rules of Evidence, a section of the Columbia Product Liability Act, two U.S. Supreme Court Cases, and one from the 15th Circuit.**[8]

Does this look familiar? The structure and the call of the question make it appear to be a "closed packet" office memorandum of law, the kind of assignment that forms the core of the traditional first-semester LRWR course. But guess again, because this is an excerpt from a "Performance Test" prepared by the National Conference of Bar Examiners (NCBE) and included in the bar exams of several states in July 1993.[9] Thus, an evaluative and educational tool that is common to most legal writing programs in the United States[10] is now being used as part of a

---

[8] Jane Peterson Smith, *The July 1993 Performance Test Research Project*, 64 THE BAR EXAMINER 36, 38 (May 1995). (See discussion in note 9, *infra*, and accompanying text.)

[9] *Id.* As part of the project, 1700 applicants in three states that had not previously used performance tests (Georgia, New Mexico, and Virginia) were given at least one of these NCBE performance tests, and an even greater number of applicants in three other states that *had* previously used performance tests (Alaska, Colorado, and California) were given at least one of these NCBE tests. *Id.* at 40. The tests were drafted for the NCBE by "highly respected clinical legal educators who [had] ten years' experience in creating tests for the California bar examiners." *Id.* at 37. The answers were scored by six teams of graders from ten different states. *Id.* The test consisted of a fact summary, a "file" with additional facts set forth on various documents, and a "library" of authorities. *Id.* at 38-39.

[10] The author of a recent study has concluded that only half of the programs in the country use "closed packet" assignments. Ramsfield, *supra* note 3, at 5, 76. However, the

limited-time exam by the bar examiners of several states to measure the very skills under consideration here.[11] California has used one full day of "Performance Tests" in each test administration since 1983.[12] Alaska has been using a half-day Research Question since July 1982.[13] Colorado has been using some similar questions since February 1989.[14] In 1996, the NCBE gave notice to states that it would begin offering two Multistate Performance Test questions for each test administration, and four states have opted to use at least one of the questions in their February 1997 administrations.[15] In 1996, the director of testing for the NCBE estimated that "up to half of the states [would] sign on within the next three to five years."[16]

---

actual question upon which this conclusion is based asks "What research assignments are required in the first-year LRW course?," *id.* at 53, and some respondents might have failed to list "closed packet research" in answer to this question on grounds that "closed packet" assignments are not, technically speaking, research assignments at all.

[11] A recent article on these new kind of questions explained how they differ from other questions:

> In contrast [to essay or multiple-choice questions, these "Performance Test" or] PT problems ask candidates to carry out the kinds of activities lawyers actually perform, such as drafting a brief in support of a motion. PT problems require applicants to read and analyze a wide array of documents with which attorneys normally work (such as statutes, regulations, legal opinions, transcripts of depositions, briefs, police reports, news articles, interview notes, investigator reports, and internal memos). Applicants must sift through several pages of these materials to identify the information that is salient to their assigned task and then prioritize and organize that information in preparing their response. Unlike the typical MBE or essay question, a PT problem includes documents that contain the specific laws that are needed to respond. In this sense, PT problems are more like an "open book" exam. As in practice, information sources for PT problems (such as witness accounts of events) may be unreliable or biased; facts are sometimes ambiguous, incomplete, or even conflicting; and applicants have to recognize these problems and determine how to cope with them.

Stephen P. Klein, *The Costs and Benefits of Performance Testing on the Bar Examination*, 65 THE BAR EXAMINER 13 (Aug. 1996).

[12] *See* Jane Peterson Smith, *Performance Testing in California*, 1983-1989, 58 THE BAR EXAMINER 17 (Aug. 1989); Jane Peterson Smith, *Reporting and Grading the California Performance Test*, 53 THE BAR EXAMINER 16 (Feb. 1984).

[13] Jeffrey M. Feldman & Margie MacNeille, *Certifying Professional Competence: The Alaska Experiment*, 52 THE BAR EXAMINER 4 (Feb. 1983).

[14] Alan Ogden, *Performance Testing in Colorado*, 58 THE BAR EXAMINER 19 (Nov. 1989).

[15] Conversation with NCBE administrator Nancy Holder, January 22, 1997. The Texas Board of Law Examiners recently announced that it would begin using one of the MPT questions in its February 1998 administration. Letter from Rachael Martin, Executive Director, January 29, 1997.

[16] Brendan Kirby, *The Bar Exam Is Changing*, NATIONAL JURIST 26 (Feb./March 1996). The introduction of the performance test in California was controversial in some respects, but bar exams have always been subject to criticism on grounds that they fail to test "real-life" skills, so it strikes me as harsh to criticize California for trying to deal

If the question above was not recognizable, how about the skills it purports to measure? Here is an outline of the skills the NCBE says its performance tests are designed to measure:[17]

> **A. Legal Analysis: The ability to analyze and evaluate legal authority, identify legal issues, and generate, assess, and justify the relative merits of alternative or competing legal positions.**
>
> **1. The ability to read and analyze cases, statutes, regulations, and other primary authorities.**
>
> **2. The ability to identify legal issues.**
>
> **3. Based on analysis of authority, analysis of facts, and identification of issues, the ability to formulate legal theories.**
>
> **4. The ability to develop and organize objective discussions or partisan arguments designed to achieve appropriate legal results.**

---

with this criticism head-on. For extended discussion of the development of performance tests, and the deficiencies they purport to correct, see D. L. Alderman, F.R. Evans & G.Wilder, *The Validity of Written Simulation Exercises for Assessing Clinical Skills in Legal Education*, 41 EDUC. & PSYCH. MEASUREMENT 1115 (1981); David R. Brink, *The Testing and Admission Process in the Year 2000*, 54 THE BAR EXAMINER 10 (Feb. 1985); S.B. Dunbar, D.M. Koretz & H.D. Hoover, *Quality Control in the Development and Use of Performance Assessments*, 4 APPLIED MEASUREMENT IN EDUC. 289 (1987); Daniel R. Hansen, *Do We Need the Bar Examination? A Critical Evaluation of the Justifications for the Bar Examination and Proposed Alternatives*, 45 CASE W. RES. L. REV. 1191 (1995); Joseph D. Harbaugh, *Examining Lawyers' Skills*, 59 THE BAR EXAMINER 9 (Nov. 1990); R.C. Linn, E.L. Baker & S.B. Dunbar, *Complex, Performance-Based Assessment: Expectations and Validation Criteria*, 20 EDUC. RESEARCHER 15 (1991); Marcia Kuechenmeister, *A Performance Test of Lawyering Skills: A Study of Content Validity*, 64 THE BAR EXAMINER 23 (May 1995); Stephen P. Klein, *The Costs and Benefits of Performance Testing on the Bar Examination*, 65 THE BAR EXAMINER 13 (Aug. 1996); Charles S. Kunce & Scott E. Arbet, *A Performance Test of Lawyering Skills: Candidate Perceptions*, 64 THE BAR EXAMINER 43 (May 1995); John F. O'Hara & Stephen P. Klein, *Is the Bar Examination an Adequate Measure of Lawyer Competence?* 50 THE BAR EXAMINER 28 (Aug. 1981); Douglas D. Roche, Armando M. Menocal, III, Jane Peterson Smith & Albert M. Sacks, *Performance Testing: A Valuable New Dimension or a Waste of Time and Money?* 52 THE BAR EXAMINER 12 (Nov. 1983); Rachel Slaughter, Charles Kunce & Marcia Kuechenmeister, *Bar Examinations: Performance or Multiple Choice?* 63 THE BAR EXAMINER 7 (Aug. 1994); Jane Peterson Smith, *MPT Update*, 64 THE BAR EXAMINER 28 (Aug. 1995).

[17] This outline is derived from the NCBE's own test specifications; I have omitted the sub-subheadings in the interest of saving space. Marcia A. Kuechenmeister, *A Performance Test of Lawyering Skills: A Study of Content Validity*, 64 THE BAR EXAMINER 23, 31-33 (May 1995).

5. The ability to evaluate legal theories and arguments in order to predict outcome of authoritative decision maker.

B. Fact Analysis: Given knowledge of legal rules and preliminary identification of legal issues, the ability to identify areas of factual inquiry, assess facts and marshal facts in support of legal arguments.

1. The ability to identify relevant facts.

2. The ability to acquire additional relevant facts.

3. The ability to identify inconsistencies among facts.

4. The ability to identify the reliability of asserted facts.

5. The ability to group and categorize facts in terms of legal rules.

6. The ability to select aspects of the facts which appear to call for the application or non-application of a legal rule or concept.

C. Problem-Solving: The ability to identify and diagnose problems in terms of client objectives and to suggest measures to achieve those client objectives.

1. The ability to isolate the client's problem and to identify, generate, and organize factual and legal information in a way that facilitates the formation of alternative solutions.

2. The ability to develop and evaluate alternative courses of action designed to advance some or all of the client's objectives.

3. The ability to implement the selected solution by adopting strategies and tactics and by taking action to advance the client's objectives.[18]

---

[18] Before making the decision to make performance tests available to interested states, the NCBE put together four prototype performance tests and a set of test specifications. *Id.* at 23. The NCBE then submitted both the tests and the specifications to what it called a "Content Validity Study Panel" for a one-day evaluation session. *Id.* at 24. The panel agreed that the skills listed are "fundamental," and apparently agreed

Given that the bar examiners are reaching the conclusion that most of the skills we are concerned about can be tested, are we bound to reach the same conclusion, or can we draw some distinctions between their purposes and our own? One distinction might be that bar examinations are designed to separate those candidates who are minimally qualified to begin the practice of law from those candidates who are not minimally qualified to begin the practice of law, while an examination in a legal writing course (at least, a graded one) should be designed in part to establish relative rankings of student performance. Yet this distinction probably cannot withstand scrutiny. After all, no state uses a performance test to the exclusion of other kinds of testing, and on most bar examinations, a strong showing on one portion of the examination can counteract the effects of a poor showing on another portion. This means that, in effect, the performance portion is being used to rank candidates for admission to the bar just as much as a similar examination would be used in a writing course to rank students.

Are there other skills that simply cannot be tested in a limited-time setting? What about research? I have already discussed one option for measuring legal research skills: the limited-time commando-style library raid. But students might also be questioned more directly about research, as in the questions that follow:

19. Name the two national legal encyclopedias.

20. May headnotes appearing in the National Reporter System be cited as correct statements of the law? Why or why not?

21. Give the name of the CD-ROM index to legal periodicals that we have in our law school library, and explain the chief advantages of using it rather than its print counterpart.

---

that the drafters of the specifications had properly focused on the concerns expressed in a recently-issued ABA report on legal education. *Id.* at 25 (citing American Bar Association, LEGAL EDUCATION AND PROFESSIONAL DEVELOPMENT — AN EDUCATIONAL CONTINUUM, Report of The Task Force on Law Schools and the Profession: Narrowing the Gap (1992) (hereinafter "MacCrate Report")); *see also* Jane Peterson Smith, *supra* note 8, at 37 (the specifications "were developed using the California model and the new MacCrate list of fundamental skills"). Interestingly, the ABA task force itself cited the increased use of performance tests on bar examinations as a positive development, although it noted that several criticisms had been leveled at the performance tests used up to that time. MacCrate Report at 280-282.

What about writing style or *Bluebook* form? It might be objected that these skills cannot be tested apart from a large project. But it is not so difficult to come up with questions that get at these skills. Many programs give some form of *Bluebook* examination. As for style, a student might be asked to rewrite a piece of poor writing, or to choose the best and worst from a series of selections. If these questions are written with care, they will operate to measure most of what we are looking for in large writing projects.

So what is left? Again, if there is some intangible skill that can be measured only by allowing a student to ruminate, sift, cogitate, outline, draft, redraft, and polish over a series of weeks (and I am perfectly willing to concede that there is), a limited-time exam cannot, by definition, measure that skill. But this does not end the debate, because, as stated earlier, this skill might be measured by another assignment, either in the same course or in a different course. In addition to the option of giving both a graded examination and a large graded writing project in the same course, schools have the option of waiting to measure the more "intangible" skill in a later semester. For example, a three-semester LRWR program might use an examination as the sole means of evaluation in the first semester, but use an objective memorandum of law in the second semester, and an appellate brief in the third semester.

IV. FOR MOST PROGRAMS, THE ADVANTAGES OF USING EXAMINATIONS OUTWEIGH THE DISADVANTAGES

Here is the most controversial proposition. If you have not used an examination in the past, the prospect of constructing, administering, and grading an examination might be a daunting one. And for some programs, an examination may not be feasible. But before you reject the idea, take a long look at the perceived difficulties and disadvantages.

The single largest disadvantage or cost incurred as a result of using an examination is obvious: examinations take additional time to prepare and grade. There is no avoiding this proposition. There have been weeks in which I wondered why we ever decided to give examinations. But preparing and grading examinations may not take as much time as you think. One reason for this is that there are many questions that can be used every time the course is offered: many of the principles being tested are simply fundamental, and there are not that many different ways of asking about them. A technique that fits in with

this approach is to give students a fairly large list of "study questions" at the beginning of the semester, announcing at the same time that a certain percentage of the questions on the final examination will be drawn from these study questions. One of my colleagues has used this technique with great success.

A related approach is to have a departmental pool of examination questions, which also cuts down on the time needed for examination preparation. Courses need not be cookie-cutter copies in order to use some of the same questions on the examination. Departments might even employ a uniform examination, which means that the preparation burden on each individual instructor is minimal.

As for grading, the time spent here can be reduced substantially, depending upon how much of the material you decide to cover with questions that can be graded electronically (usually multiple choice). It may seem counterintuitive to use multiple-choice questions in a course that has the word "writing" in the title, but once you come to the realization that most of the skills we teach can be broken down into their constituent parts, I think it follows that the multiple-choice format will work for many of these parts.[19] The validity of multiple-choice testing in the law field has been established by fairly careful study,[20] and

---

[19] After years of using mostly short-answer questions, every Assistant Professor in the legal writing department at South Texas College of Law switched to all multiple-choice questions in fall 1996. The cost in time spent drafting questions in September, October, and November was more than repaid by the reduction in grading time in December.

[20] *See* references listed in Rachel Adams, *Multiple-choice Item Writing: Art and Science*, 61 THE BAR EXAMINER 5, 14 (Feb. 1992). Adams describes the benefits of multiple-choice testing as follows:

> First of all, the multiple-choice format permits the assessment of a large body of content knowledge in a relatively short time. Second, a test consisting of many well-constructed multiple-choice questions which appropriately sample the knowledge domain can produce highly stable test scores. Third, tests constructed of multiple-choice questions can be efficiently, objectively, and accurately scored. Fourth, well-constructed multiple-choice questions can demonstrate an examinee's ability to apply a broad body of knowledge, as well as the ability to remember facts and principles. Finally, reliable scores from well-constructed multiple-choice tests can be combined with scores from performance tests to more adequately demonstrate both an examinee's grasp of the large body of content knowledge necessary for good professional practice and an examinee's skills in specific or delimited areas of expertise such as writing an analysis of a case.

Id. at 5. For further discussion of the benefits of multiple-choice testing, see Steven J. Osterlind, CONSTRUCTING TEST ITEMS 175-76 (1989); J. Marshall Trieber, *The Use of Multiple-Choice for Testing*, 34 TRAINING & DEV. J. 24 (Oct. 1980); John A. Winterbottom, Use of Essay and Objective Techniques in Bar Examinations, Address at the National Conference of Bar Examiners 1968 Conference *in* 38 THE BAR EXAMINER 5 (Jul. 1969).

there appears to be fairly widespread acceptance of this form of testing not only on bar examinations, but also in law school examinations.[21]

Admittedly, this reduction in grading time does come at the expense of additional preparation time. Good multiple-choice questions don't drop out of the sky: they can be devilishly difficult to construct and fine tune. Here are some suggestions, gleaned from some of the literature on educational testing and my own experience.

First, get as many people to look at the question as you can, within the limits imposed by security concerns. The most innocent-looking sentence can sometimes hide the most awful ambiguity. Be particularly careful to avoid double negatives.

Second, try to make the number of questions on each area roughly proportionate to the amount of attention paid to that area during the semester.

Third, Consider deductions for wrong answers. The NCBE does not employ this method of scoring on the Multistate Bar Examination,[22] nor does the Law School Admission Council (LSAC) employ it on the LSAT,[23] but the literature suggests that this method can yield a more accurate curve, because it penalizes students who guess from among several choices more than it penalizes students who guess from among just two choices. In other words, statistically speaking, for the set of students who guess on at least one item, there will be a greater difference between the scores of, on the one hand, students who

---

[21] Steve H. Nickles, *Examining and Grading in American Law Schools*, 30 ARK. L. REV. 411, 434, 447-51 (1977). Some legal educators and members of the bar no doubt are concerned about the increased use of machine-scored questions on law school examinations, insofar as they view essay questions as testing different competencies. (As Nickles points out, the idea that essay examinations have "superior power to expose organizational, integrational, and synthesizing abilities of students, which are the abilities to be assessed above all others" in law school is "so firmly embedded as to give the appearance of a truism." *Id.* at 447-47.) It seems unlikely that increased use of multiple-choice questions will have serious long-term consequences, so long as a substantial portion of law school examinations use at least one essay question. To the extent that a law school depends upon LRWR courses to ensure development of the writing skills, and to the extent that use of a multiple-choice examination in a LRWR course means less emphasis on writing in the course, there might be some legitimate concern here. However, given the number of essay examinations still being given in law schools, and the fact that most LRWR courses will still be giving great emphasis to these writing skills, there would be little danger arising from use of multiple-choice questions on our examinations at this time.

[22] *See, e.g.,* 1984 Multistate Bar Examination Information Booklet 1 ("[s]cores will be based on the number of questions answered correctly").

[23] *See* 1996-1997 LSAT/LSDAS Registration and Information Book 92.

are able, because of superior knowledge, to eliminate at least one of the choices, and, on the other hand, students who are able to eliminate none of the choices. The formula to use for deductions is 1/X-1, where X equals the number of answers available. Thus, an incorrect answer on a question with four choices should result in a deduction of ⅓ of a point.[24]

Fourth, if you deduct points for incorrect answers, it is very important not to add any answers that are obviously incorrect, because the penalty will then be diluted.[25] Thus, in constructing a question, if you can come up with only two plausible incorrect answers, it is best to simply give the student three choices, rather than adding a fourth, even if most of the other questions on the exam have four choices.[26] (But if you are not deducting points for incorrect answers, the absolute number of choices is not important: it is statistically important only that you have at least two plausible incorrect answers.[27])

Fifth, avoid questions with a tri-level structure, if possible. Tri-level questions are structured as follows: the top level contains the fact pattern or nucleus of facts; the middle level contains a series of conclusions; and the bottom level contains a series of statements about which of the *middle-level* conclusions are correct.[28] It is difficult to write these questions so that good

---

[24] Howard J. Gensler, *Valid Objective Test Construction*, 60 ST. JOHN'S L. REV. 288, 291-93 (1986).

[25] *Id.* at 293.

[26] *Id.* at 291, 293.

[27] *Id.* at 290-91.

[28] Howard J. Gensler, *Statistical Instability of the Tripartite Multiple Choice Item on the Multistate Bar Exam*, 14 GLENDALE L. REV. 25 (1995). Here is an example Gensler uses:

> Tom was in a car accident and was not expected to live through the night. He asked his son, Sam, who was a lawyer, to make his will. Tom left half of his estate to his daughter Ann, one tenth to Sam, and the rest to his church. Tom was lucid at the time, and specifically disinherited his wife, Wilma. Tom executed the will which Sam had written by hand, and it was witnessed by three nurses. It was not dated. Tom died that night.
>
> I. The will is invalid because Tom, a beneficiary, drew up the will.
> II. The will is invalid because Wilma was disinherited.
> III. The will is invalid because it was handwritten and it did not conform to the requirements of a holographic will.
>
> A. I, II, and III.
> B. I and II only.
> C. II only.
> D. None.

Id. at 26.

students get them right and poor students get them wrong.[29]

Sixth, try to make sure that the questions on the examination cover a range of difficulty. (This provides a nice check on the examination as a whole, as explained in the next numbered suggestion.) For example, I have often used a fairly simple "matching" question, in part to make the students feel a little more confident about the examination while they are taking it, and in part as a kind of check on the exam itself. (See Appendix.) This matching question could be easily converted to multiple choice.

Seventh, consider doing a post-exam statistical analysis of each question to determine the degree of correlation between the group of students who performed well overall on the exam and the group of students who performed well on each particular question.[30] This will give you a sense of which questions worked well. (A question worked well when the students who performed well on the entire exam tended to do well on that question, and the students who did poorly overall did poorly on that question.) Perhaps even more importantly, this analysis will point up any questions that turned out to be disasters. (If, as to a particular question, the students performing well on the exam overall did no better, or did worse, than the students who performed poorly, you can say with some assurance that there was a problem with the question.) You may even want to strike such questions from consideration when you are scoring the exam.[31]

Another perceived disadvantage is that an examination shifts the focus of the course away from the writing of documents. One of the strengths of our courses is the close relation

---

[29] Gensler, *supra* note 24, at 294. There is some controversy about this analysis of tri-level questions. Apparently, some of the drafters at the NCBE agree that this structure is flawed, yet the MBE continues to be offered with these items included, which has led to even more criticism. Gensler, *supra* note 28, at 25, citing Adams, *supra* note 20, at 10. So far as I know, the Educational Testing Service (ETS) continues to use them on the Scholastic Achievement Test (SAT) and the Law School Admission Council (LSAC) continues to use them on the Law School Admission Test (LSAT). I see some possible problems with Professor Gensler's analysis, and thus am not absolutely convinced that this form of question is fundamentally flawed, but my lack of familiarity with statistical analyses (and my own difficulties in drafting satisfactory tri-level questions) makes me hesitant to ignore his warning.

[30] Howard Gensler, *supra* note 24, at 294-96.

[31] *Id.* at 296. I have not done this, although I agree with the principle. I still have a gut reaction that it would be somehow unfair to strike the question entirely, and if I have this reaction, most assuredly some of the students who "correctly" answered the question would have the same reaction, and the energy spent dealing with these students would be better spent elsewhere. As you can see, I'm still struggling with this.

between the main projects we require students to complete and the kinds of projects they will ultimately do in practice: after all, at their first law jobs, our students won't be answering multiple choice questions all day long. But I think it is a mistake to exclude, on this basis, all other kinds of projects. Examinations can complement other projects just as much as they can distract from other projects. For example, a professor could give a fairly straightforward "closed packet" memorandum problem as a three-hour examination in mid-semester, and then assign the same facts as a long-term open memorandum project a few days later, so that students can build on the authorities and the analysis they have already worked with. Indeed, one might argue that there is more realism in a three-hour "performance" examination than in a six-week memorandum project, given the short turnaround time demanded by some law firm supervisors.

Of course, another consequence or aspect of shifting the focus of the course is that students have less time and energy to devote to the more traditional writing projects. This objection is particularly valid if the examination is offered during the regular examination period, not long after the large project of the semester has been completed. In expecting students to finish the large writing project, then turn around and prepare for all their other examinations, as well as preparing for an examination in their writing courses, we are expecting a great deal.[32]

There would seem to be at least a couple of ways of ameliorating this difficulty. One would be to scale back on the major project, or make it less difficult. Another would be to administer the examination earlier in the semester. After all, most of what we expect students to know in order to complete the large project has already been given to them before they begin that project. Thus, most of the questions on the final examination I give in the first-semester course could be answered by the students at the end of the ninth week of classes, and I'm sometimes tempted, when I'm planning the semester, to schedule the examination for that week. (Of course, this would normally require the permission of the associate dean for academic affairs, or

---

[32] Surprisingly, at South Texas I hear very little grumbling about the LRWR examination, which is almost always given during the regular examination period. I am grateful that the examination was already in place when I came to South Texas; imposing an examination on students who knew that it had not been required in past semesters would probably generate more discontent.

some faculty committee, or both, which itself could be viewed as a disadvantage.)

In fairness, I should add that the one time that our department gave the examination in the middle of the semester, the students were not happy. And they expressed their displeasure not so much at the time of the examination, but at the end of the semester: many claimed that the material covered in the first seven or eight weeks did not "gel" for them until they were nearing completion of the large open memorandum of law, so that the examination was not a fair measurement. (This was particularly true of the research skills covered in the examination, according to the students.) We did not feel particularly guilty as a result of doing this, reasoning that those students for whom things "gelled" later would presumably receive higher grades on their memos, and would not be significantly penalized, but we did switch back to offering the examination during the regular examination period at the end of the semester.

Against these disadvantages or costs must be set the advantages or benefits of using an examination. To the advantages listed in Section II, above, we can add the following concrete benefits.

First, you can evaluate knowledge of principles that are important, but that you were unable to incorporate into the large-scale writing project that semester, or that you find difficult to evaluate independently when grading the large-scale project. For example, there may be some idiosyncracies in how statutory research must be conducted in your state. If you have only one large graded project, and the problem you choose implicates only federal law or the law of some other state, competency or lack of competency in state statutory research will not be reflected in the grade. Texas happens to have several small idiosyncracies in this area, but I can measure competency independently here with a question like this:

> **22. The senior partner says that in 1965 the firm represented a client in an unusual situation governed by a Texas statute, and the partner still has the memorandum written in that case. An almost identical situation has arisen, but the information in the memorandum written for the 1965 problem is too old to be reliable. How would you update the memorandum?**

(To make a question like this even more realistic, you might even provide some of the facts and some of the relevant statu-

tory provisions and appropriate tables and supplements. The Lawyering Skills Program at Albany Law School employs this technique, which seems like it might work extremely well.)

Second, you can increase incentives for your students to work hard on assignments that do not themselves *directly* affect the course grade, but which are more obviously related to the course exam than they are to the large writing project. (This is really a corollary of the first advantage: when students can see more clearly that an assignment will probably have some effect on their course grades, they tend to work harder on that assignment.)

Third, you can evaluate knowledge and skills without much fear that students are collaborating improperly or cheating in some other way.[33]

Fourth, you can neutralize to some degree the natural advantage held by students who are skilled writers when they arrive at law school. (I recognize that some of you will not agree that this is a benefit, either because you do not agree that skilled writers hold an advantage, or because you believe there is no need to neutralize the advantage. However, remember that perception or feeling can sometimes be as important as reality. Many students walk into our courses feeling unfairly disavantaged because they are not strong writers, and some walk out feeling the same way.)

Those of us who are making careers in the field of legal writing have sometimes been frustrated with our non-legal-writing colleagues' refusal to at least *consider* curricular change. The argument that "we can't do that because we've never done it before" simply doesn't carry much independent weight, regardless of who makes it. Consider using an exam.

---

[33] From discussions over the years, on Legwri-l (the Legal Writing listserve) and elsewhere, I know that some of us view this as a more significant advantage than others of us do.

APPENDIX

MATCH THESE . . .	TO THESE
___ American Law Reports (ALR)	A. An individual pamphlet issued upon enactment of a statute into law.
___ Digest	B. The same case in a different reporter.
___ Law Reviews or Law Journals	C. A system of arranging, by subject, the headnotes of cases.
___ Official Reports	D. Statutes enacted by a legislature, arranged chronologically.
___ Parallel Citation	E. The first appearance of a decision issued by a court.
___ Restatements	F. Alphabetical compilation of terms defined by the courts.
___ Session Laws	G. The public, general, and permanent statutes of a
___ Shepard's Citators	jurisdiction in a fixed subject or topical arrangement.
___ Slip Law	H. The opinions of a given court published by the
___ Slip Opinion	court-designated publisher.
___ Statutory Code	I. Publisher-selected leading cases, both state and federal, with detailed annotations and
___ Treatise	commentary.
___ U.S. Statutes at Large	J. The official, permanent session law publication for federal laws.
	K. Student-edited periodicals that contain respected commentaries on the law.
	L. Specialized publications whose purpose is to trace the use of earlier authority in later sources.
	M. Respected reiteration of major legal doctrines covering ten specific fields of law.
	N. Scholarly written discourse on an area of law by an expert in the field.