

## Chapter Seven

### Best Practices for Assessing Student Learning

#### A. The Importance and Purposes of Assessments.

Grades are important in law school, particularly for first year students. After one semester, grades determine which students are eligible for Law Review, Moot Court, and other significant opportunities in law school, which students are most likely to pass the bar examination, and which students will compete for the most highly compensated jobs. Students who fare most poorly are forced to leave law school, and they lose their opportunity to become lawyers. These are high stakes.

The main purpose of assessments in educational institutions is to discover if students have achieved the learning outcomes of the course studied.<sup>681</sup> In other words, we use assessments to find out whether students are learning what we want them to learn.

In law schools, as in medical schools, one purpose of assessment is to determine which students should receive degrees, but other purposes of assessment are more important.

Aside from the need to protect the public by denying graduation to those few trainees who are not expected to overcome their deficiencies, the outcomes of assessment should be to foster learning, inspire confidence in the learner, enhance the learner's ability to self-monitor, and drive institutional self-assessment and curricular change.<sup>682</sup>

An institution's decisions about what and how it assesses student learning reflect the values of the institution.

Assessment is also a statement of institutional values. Devoting valuable curricular time to peer assessment of professionalism, for example, can promote those values that are assessed while encouraging curricular coherence and faculty development, especially if there are corresponding efforts at the institution toward self-assessment and change.<sup>683</sup>

The goals and methods we select for assessment directly affect student learning. "Assessment methods and requirements probably have a greater influence on how and what students learn than any other single factor. This influence may well be of greater importance than the impact of teaching materials."<sup>684</sup>

---

<sup>681</sup> ALISON BONE, NATIONAL CENTRE FOR LEGAL EDUCATION, ENSURING SUCCESSFUL ASSESSMENT 3 (Roger Burrige & Tracey Varnava eds., 1999), available at <http://www.ukcle.ac.uk/resources/assessment/bone.pdf> (last visited April 27, 2006).

<sup>682</sup> Epstein & Hundert, *supra* note 150, at 226.

<sup>683</sup> *Id.* at 231.

<sup>684</sup> BONE, *supra* note 681, at 2.

[C]hanging the assessment procedure is one of the most effective ways of changing how and what students learn. Surface approaches are induced by excessive workloads, a narrow band of assessment techniques and undue emphasis upon knowledge reproduction. Deep approaches are influenced by choice, a variety of assessment methods, project work and an emphasis upon tasks that demand demonstration of understanding.<sup>685</sup>

Thus, legal educators should consider carefully what we are trying to assess and how we are doing it.

## B. The Shortcomings of Current Assessment Practices in the United States.

In the traditional law school course, especially in the all important first year, the only evaluation of how well a student is learning, and the entire basis for the student's grade for the course, is a three hour end-of-the-semester essay exam that requires students to apply memorized legal principles to hypothetical fact patterns. The practice of basing the assessment of student learning on a single test was initiated in the early 1870's at Harvard Law School by Dean Christopher Langdell.<sup>686</sup> Prior to that date, other American law schools relied upon frequent oral quizzes, evaluation of moot court performances, and, in jurisdictions that accorded graduates of local law schools diploma privileges, comprehensive written examinations requiring descriptive essays on relevant points of law.<sup>687</sup> American law schools quickly copied the Harvard way and "by the end of the nineteenth century, the use of single exams to assess student performance had become widespread among American law schools."<sup>688</sup> The single exam tradition remains with us today, despite long-standing criticisms from academics, practitioners, and students.<sup>689</sup>

The primary reason to administer assessments is to find out whether our students are learning what we want them to learn. Judith Wegner's study of legal education determined that the current grading practices of legal educators in the United States function less as a means for measuring student learning than as a means for sorting and ranking students and for "weeding out" students who are not

<sup>685</sup> *Id.* at 4.

<sup>686</sup> Aizen, *supra* note 313, at 768-69 (citing Steven Friedland, *A Critical Inquiry into the Traditional Uses of Law School Evaluation*, 23 PACE L. REV. 147 (2002); Steve Sheppard, *An Informal History of How Law Schools Evaluate Students, With a Predictable Emphasis on Law School Final Exams*, 65 UMKC L. REV. 657 (1997); John J. Costonis, *The MacCrate Report: Of Loaves, Fishes, and the Future of American Legal Education*, 43 J. LEGAL EDUC. 157 (1993); Russell L. Weaver, *Langdell's Legacy: Living with the Case Method*, 36 VILL. L. REV. 517 (1991)).

<sup>687</sup> Sheppard, *supra* note 686.

<sup>688</sup> Aizen, *supra* note 313, at 768-69.

<sup>689</sup> For a collection of scholarship documenting the dissatisfaction with the single exam practice and supporting an increase in the number, variety, and quality of law school assessments, see *id.* at 769 nn.19 & 20. Recommendation 6 of the ABA's Task Force on Lawyer Competency was that "[l]aw schools and law teachers should develop and use more comprehensive methods of measuring law student performance than the typical end-of-the-term examination. Students should be given detailed critiques of their performance." CRAMTON REPORT, *supra* note 275, at 4.

developing the requisite knowledge, skills, and values to pass a bar examination.<sup>690</sup> She concluded that our emphasis on using assessments as a sorting device impedes the effectiveness of our educational efforts. “Since the point of law school is to foster learning and to develop learning habits such as professionals need, the cost of confounding learning in order to engage in incessant sorting seems very large indeed.”<sup>691</sup>

The Carnegie Foundation’s study of legal education discovered that current first year assessment practices have harmful effects on students’ motivation and opinions of law school.

Students’ comments about assessment in their first year of law school often expressed puzzlement, frustration, and anguish. A recurring theme in their comments, striking in its frequency, was that they were not being tested on what they studied for and what they knew. Many felt that the testing was unfair, counterproductive, demoralizing, and arbitrary. Students saw little or no relation between their classroom experience and the end-of-the-semester examinations, or between learning to be a good lawyer and doing well on exams – a criticism that has been leveled at the cognitive apprenticeship in many professional and graduate schools. As our earlier chapters showed, law schools’ heavy emphasis upon academic training, in contrast to the education in settings of practice typical of preparation for the health professions, heightens the likelihood of a disparity between learning to be a law student and learning to be a lawyer.

A number of students complained that the quality and quantity of their studying was unrelated to their performance on the final examination. They claim to have had little feedback during the semester and no basis on which to gauge whether they were mastering the material or making adequate progress toward the desired proficiencies.<sup>692</sup>

The scaled grading system allows law schools to sort students for legal employers, but it impedes learning, community building, and moral development.

The current scaled grading system in most law schools, which is based solely upon comparison to and competition with other students, is not a system designed to promote either community or the broader ideal of justice. It is a prime example of the hierarchical systems Mary Rose O’Reilly places on moral notice . . . . It is entirely individual-focused and rights based. It is judgmental and exclusive rather than compassionate and inclusive. It is essentially designed to rank students in an important but limited area of legal skills (while ignoring other important indices of qualifications as a lawyer) for the convenience of firms who are in the job market. If it has a

---

<sup>690</sup> Wegner, *Assessment*, *supra* note 24, at 19-22 and 34. We are only somewhat successful in preparing students for bar examinations, given that only 50% to 80% of law school graduates pass a state bar examination on their first attempt.

<sup>691</sup> *Id.* at 33.

<sup>692</sup> SULLIVAN ET AL., *supra* note 7, at 206.

pedagogical purpose, it is only to spur students to study for grades in competition with their fellow students, a “benefit” which is lost on many students after the first year when they see where they stand in the class and give up on trying to rise any higher. The competitive grading system is a primary instrument separating students from faculty in law schools and separating students from other students. It is a central impediment to construction of an effective law school community.<sup>693</sup>

Despite its long history as a part of legal education, the end-of-the-semester essay exam is an inadequate method for assessing student learning, and fundamental aspects of our current practice are significantly flawed. As Sandy D’Alemberte put it, “Is there any educational theorist who would endorse a program that has students take a class for a full semester or a full year and get a single examination at the end? People who conduct that kind of educational program are not trying to educate.”<sup>694</sup>

As currently used, the end-of-the-semester essay exam is neither valid, nor reliable, nor fair. The problems with our current practice were summarized by Judith Wegner as follows:

In sum, the current assessment system has a number of significant costs worth reconsidering: compromised efficacy that results from conflating sorting students and evaluating learning; perpetuation of past advantages and disadvantages in unintentional ways; confusion that impedes learning; and deployment of faculty time in relatively ineffective ways. In light of these costs, it is worth endeavoring to develop new systems of assessment deliberately designed to foster learning.<sup>695</sup>

Most of the preceding comments relate to assessment practices in traditional doctrinal courses. Unfortunately, current assessment practices are also flawed in experiential education courses such as simulation-based courses, in-house clinics, and externships.

In simulation-based courses, the primary and sometimes sole method of assessment is for a single teacher to observe a student performing a limited number of lawyering tasks. Sometimes, self- or peer-evaluation is also used. Frequently, students are given a grade on every performance, often without any opportunity to receive formative feedback before the summative assessment and without any opportunity to continue practicing until the appropriate level of proficiency is achieved. For that matter, almost no effort has been made to describe appropriate levels of proficiency.

In many in-house clinics and externships, grades are based mostly on the subjective opinion of one teacher who supervises the students’ work. Grades in these courses tend to reflect an appraisal of students’ overall performance as lawyers, not necessarily what they learned or how their abilities developed during the course. When written criteria are given to students, they tend to be checklists that cover the

---

<sup>693</sup> BENNETT, *supra* note 70, at 170.

<sup>694</sup> D’Alemberte, *supra* note 14, at 52.

<sup>695</sup> Wegner, *Assessment*, *supra* note 24, at 33.

entire spectrum of lawyering activities without any descriptions of different levels of proficiency.<sup>696</sup>

Virtually no experiential education courses give written tests or otherwise try to find out if students are acquiring the knowledge and understandings that the courses purport to teach. Items that could be clearly subjected to more objective testing include students' understanding of theories of practice or particular aspects of law, procedure, ethics, and professionalism. A students' understanding of many aspects of law practice as well as his or her lifelong learning skills could also be assessed, for example, by asking students to analyze recordings or transcripts of lawyers' performances. Serious efforts to assess student learning in experiential learning courses are not being made on any large scale.

In sum, except perhaps in legal writing and research courses, the current assessment practices used by most law teachers are abominable. We share Judith Wegner's conclusion that "[a] better assessment system would find ways to stimulate student reflection on future professional paths, strengths and weaknesses and guide students toward relevant learning opportunities; provide incentives that lead students to take more active responsibility for their own learning as they undertake increasingly sophisticated work throughout students' law school careers; and document information that would attest to graduates' professional capabilities while assisting employers in making efficient and informed hiring decisions."<sup>697</sup>

Legal educators in the United States "need to clarify the purposes of grading systems, reconsider practices that perpetuate advantages and disadvantages associated with high-stakes testing early in students' law school careers, find ways to stimulate rather than skew student learning and reallocate faculty time spent on semester-end grading to better use."<sup>698</sup>

### C. Best Practices for Assessing Student Learning.

Effective assessment exhibits qualities of validity, reliability, and fairness.<sup>699</sup> Validity means that an assessment tool must accomplish the purpose for which it was intended. Reliability means the test or measuring procedure yields the same results on repeated trials. A single do-or-die final essay exam given under time pressure at the end of the semester fails all three criteria.<sup>700</sup> It is neither valid, nor reliable, nor fair.

The best practices described in this section reflect recommendations for improving assessment practices arising from the work of numerous scholars, including Judith Wegner's study of legal education for the Carnegie Foundation for the Advancement of Teaching. They incorporate the five key principles that Wegner believes should influence the design process of an improved assessment system:

---

<sup>696</sup> See, e.g., Appendix A and B in Stacy L. Brustin & David F. Chavkin, *Testing the Grades: Evaluating Grading Models in Clinical Legal Education*, 3 CLINICAL L. REV. 299 (1997).

<sup>697</sup> Wegner, Assessment, *supra* note 24, at 63.

<sup>698</sup> *Id.* at 30.

<sup>699</sup> MICHAEL JOSEPHSON, LEARNING AND EVALUATION IN LAW SCHOOL 7 (1984).

<sup>700</sup> Gregory S. Munro, *How Do We Know If We Are Achieving Our Goals?: Strategies for Assessing the Outcome of Curricular Innovation*, in ERASING LINES, *supra* note 38, at 229, 237.

- learning is the point,
- learning must be made visible in order to be assessed,
- learning is multifaceted and develops over time,
- assessment must reflect the particular purposes being served (such as evaluating, educating, assuring quality, conferring distinction, and documenting professional capability), and
- assessment occurs in context.<sup>701</sup>

The principles described in this section are only the beginning of the work that is needed to improve assessments in law schools. Experimentation with new methods of assessment will reveal the need to modify and add to the principles and proposals set forth below.

### 1. Be Clear About Goals of Each Assessment.

**Principle:** The teachers are clear about the goals of each assessment.

**Comments:**

It is important to know what we are trying to evaluate. The goals of a particular assessment may be to evaluate a student's knowledge, behavior (what a student does before and after a learning experience), performance (ability to perform a task), attitudes and values, or a combination of these.

*Cognitive assessment* means assessment of learning or knowledge.<sup>702</sup> For example, this could entail assessment of whether a student in Property has acquired the applicable knowledge of the substantive law. This is different from assessment of behavioral change and performance,<sup>703</sup> which is characterized by the student's ability to use knowledge.<sup>704</sup>

*Behavioral assessment* measures change in what a student does before and after a course of learning.<sup>705</sup> "This 'observation' is made concerning an event in the student's life which is not regulated, contrived, or designed for the purposes of assessment or grading."<sup>706</sup> An example would be examining whether students who studied attorney engagement agreements in their professional skills and contracts courses later recorded in the file and warned clients of the statute of limitations during their clinical internships.

*Performance assessment* measures the student's ability in a task that the student is asked to perform for the purposes of the assessment (for example, having the student find the errors in a civil complaint).<sup>707</sup>

---

<sup>701</sup> Wegner, Assessment, *supra* note 24, at 55.

<sup>702</sup> NICHOLS, *supra* note 111, at 37.

<sup>703</sup> *Id.* at 42.

<sup>704</sup> *Id.* at 37.

<sup>705</sup> *Id.* at 42.

<sup>706</sup> *Id.*

<sup>707</sup> *Id.* at 43.



*Attitudinal assessment* can measure differences in students' attitudes before and after a course of learning.<sup>708</sup> For instance, we can measure change in student attitude after a Professional Responsibility course. Law schools may want to know the attitude of incoming students on a host of issues or their perception about the law school or its programs. As student education progresses, the faculty may wish to know how particular parts of the program change student attitudes. On graduation, exit interviews may reveal attitudes the student has about her legal education, social issues, or moral issues. Finally, attitudes of practitioners toward the law school or any other relevant issues might be measured.<sup>709</sup>

## 2. Assess Whether Students Learn What is Taught (validity).

**Principle:** The assessment tools used by the teachers evaluate whether students learn what is being taught.

### Comments:

An assessment tool should be valid. An assessment tool is valid if it allows the teacher to draw inferences about the matters that the test purports to assess.<sup>710</sup> Congruence is a necessary aspect of validity, that is, the goals of the test must agree with the goals of the instruction.<sup>711</sup> For example, a professor who seeks to test students' ability to apply and distinguish cases might administer an essay question that raises issues testing the outer limits of a set of precedents. On its face, the exam appears to be a valid test of the skill. If, however, students must take the test in a closed-book setting or without sufficient time to review the relevant authorities while taking the exam, students who have developed the ability to apply and distinguish cases, but possess poor memorization skills, would likely perform poorly. Thus, the exam would not be valid.

The validity issue requires law teachers to consider carefully what law school exams measure. Referring to first year law school and other similar exams, Judith Wegner determined that "law school exams can best be understood as attempts to measure students' law-related problem-solving expertise."<sup>712</sup> Problem-based essay exams require students to perform three principle functions – spotting issues, identifying relevant authorities, and applying legal authorities to complex fact patterns – and on occasion a possible fourth, evaluating competing policies or principles.<sup>713</sup> Wegner concluded that such exams, as are typically used in the first year, "appear forthrightly directed to discerning the existence of student expertise as legal analysts confronted with a problem-solving task."<sup>714</sup>

Although essay exams appear to be a sound way to assess some aspects of problem-solving expertise, the manner in which we use them undermines their effectiveness. Law professors do not clearly explain that the purpose of the essay

<sup>708</sup> *Id.* at 44.

<sup>709</sup> MUNRO, *supra* note 4, at 115-17.

<sup>710</sup> HESS & FRIEDLAND, *supra* note 304, at 289. *See also* SMITH & RAGAN, *supra* note 201, at 95.

<sup>711</sup> SMITH & RAGAN, *surpa* note 197, at 95.

<sup>712</sup> Wegner, *Assessment*, *supra* note 24, at 3.

<sup>713</sup> Phillip Kissam, *Law School Examinations*, 42 VAND. L. REV. 433, 440-42 (1989).

<sup>714</sup> Wegner, *Assessment*, *supra* note 24, at 9.

exam is to test problem-solving expertise, and most first year courses fail to provide instruction designed to help students develop such expertise.

[C]lassroom teaching in first-year courses tends to focus primarily on certain intellectual tasks, including comprehension, analysis, application of legal principles to simple fact patterns, synthesis of related cases, and limited forms of “internal” evaluation concerning logic and doctrinal consistency. On the other hand, classes (or reading assignments) give students relatively little opportunity to observe models or experiment with application of law to complex fact patterns, synthesis across broader fields, or evaluation against the backdrop of social concerns. Students carefully observe how *others* (most notably judges) solve problems, but rarely work through how *they* (or the lawyers in key cases) might actually do so themselves. Strikingly, however, strong performance on examination essays requires demonstrated skill in just those matters that are not directly taught.<sup>715</sup>

This pattern of unintentional omission has important implications. It is extremely frustrating to some students and has a negative impact on their self-efficacy and motivation to learn.

In the view of these students, there is a significant mismatch between what professors say and do in classes and what is tested on exams. Students are not given a chance to practice what will actually be tested, and don't get feedback to gauge how they might do when the day of reckoning arrives. They don't understand how what is tested relates to what is expected of lawyers. The impression is one of enormous frustration, of effort expended to little avail, of talented learners trying their hardest, of profound puzzlement without recourse.<sup>716</sup>

The situation also gives an unfair advantage to students who have strong analytical skills when they begin law school.

Students who will be most likely to perform well under such circumstances are those who have had prior experience with (and who have internalized approaches to) similar academic tasks, those who are “expertlike” in their approaches at the same time of entry (as many faculty members probably were during their own student days), and those who have well-developed expertise in and self-awareness about learning in some other complex field that was once unknown. Others will not fare as well.<sup>717</sup>

Once the more “expert” students gain the advantage by receiving the highest grades, their expertise continues developing, and it is difficult for their peers to ever catch up.

Thus, the incongruence between what is taught and what is tested is a serious problem that legal educators should address if they want to claim that law

---

<sup>715</sup> *Id.* at 14.

<sup>716</sup> *Id.* at 6.

<sup>717</sup> *Id.* at 14.



school problem-based essay exams are fair to all students. Ensuring that law school exams test what law professors teach is an issue to consider in all courses, of course, not just those in the first year.

Before each assessment, we should consider what we expect students to learn in our courses and what is important for us to assess. Different assessment methods may be required to assess each of the following educational objectives that we might be trying to achieve:

- self-reflection and life-long learning skills,
- intellectual and analytical skills,
- core knowledge of the law,
- core understanding of the law,
- professionalism, and
- professional skills.

Our most difficult challenge, of course, is to assess the overall level of professional competence that our students possess.

### **3. Conduct Criteria-Referenced Assessments, Not Norm-Referenced (reliability).**

**Principle: The teacher conducts criteria-referenced assessments.**

**Comments:**

An assessment tool should be reliable, that is, it should accurately rate those who have learned as having learned and those who have not learned as having not learned.<sup>718</sup> It should not matter whether a student is being assessed first or last or whether one teacher or another is conducting the assessment. We join Judith Wegner and other scholars in encouraging law professors to develop and apply explicit grading criteria to minimize the risk of unreliability in assigning grades.<sup>719</sup>

Assessments can be norm-referenced or criteria-referenced. Assessments in the United States tend to be norm-referenced; assessments in the United Kingdom are typically criteria-referenced. Norm-referenced assessments are based on how students perform in relation to other students in a course rather than how well they achieve the educational objectives of the course. Normative assessment is often done to ensure that certain grade curves can be achieved.

Norm-referenced evaluations inform students how their performance relates to other students, but they do not help students understand the degree to which they achieved the educational objectives of the course. This can have a negative effect on student motivation and learning.

[S]tudents . . . perceive that something different is going on in the current circumstance, and wonder whether the “sorting” process reflects an artificial or arbitrary allocation of rewards. In the absence of a clearly stated explanation of the actual standards to be achieved,

<sup>718</sup> SMITH & RAGAN, *supra* note 197, at 97.

<sup>719</sup> See N. R. Madhava Menon, *Designing a Simulation-Based Clinical Course: Trial Advocacy*, in A HANDBOOK ON CLINICAL LEGAL EDUCATION 177, 181 (N. R. Madhava Menon ed., 1998) (“Students and evaluators need a clear understanding of the criteria on which performances will be graded.”).

it is easy to become frustrated, then angry, wasting energy that might otherwise be invested in meaningful efforts to learn.

Students also powerfully articulate their hunger to link assessment and learning. They want to learn to take exams, and they want feedback so they can improve.<sup>720</sup>

Norm-referenced assessment allows grades to be distributed along a bell curve. We should not be concerned about whether students' performances will be distributed along a normal "bell curve" because one should not expect it to be.<sup>721</sup> Mandatory grade curves are not consistent with best practices for assessing student learning. A bell curve outcome actually reflects a failure of instruction.

What matters is whether students adequately achieve the learning outcomes of the course. Our goal should be to achieve the learning outcomes we establish for our courses, whether those are to learn certain information, understand key concepts, or develop skills to a specified level of proficiency.

[T]he primary goal is to help students learn to think about their own thinking so they can use the standards of the discipline or profession to recognize shortcomings and correct their reasoning as they go. It isn't to rank students. Grading on a curve, therefore, makes no sense in this world. Students must meet certain standards of excellence, and while none of those standards may be absolute, they are not arbitrary either. Grades [should] represent clearly articulated levels of achievement.<sup>722</sup>

Some students will achieve the objectives of our courses faster or easier than other students, but if our teaching is effective and successful, all students should learn what we want them to learn and earn high marks on assessments. If a student is incapable of learning what we are trying to teach, the student should not be allowed to become a lawyer. If a student is capable of learning, but fails to do so, we may want to ask whether the fault is the student's or our own.

We can improve the quality of our assessments by following the approach used in other disciplines of developing and disclosing criteria-referenced assessments. Criteria-referenced assessments rely on detailed, explicit criteria that identify the abilities students should be demonstrating (for example, applying and distinguishing cases) and the bases on which the instructor will distinguish among excellent, good, competent, or incompetent performances.<sup>723</sup> "Ideally, criteria should be subject-based and geared specifically to the assessment to which it relates."<sup>724</sup>

The use of criteria minimizes the risk of unreliability in assigning grades.<sup>725</sup> Criteria-referenced assessment enables teachers to "judge whether certain criteria have been satisfied and normally operates on a pass/fail basis: an example would be

---

<sup>720</sup> Wegner, Assessment, *supra* note 24, at 26.

<sup>721</sup> *Id.* at 30.

<sup>722</sup> BAIN, *supra* note 299, at 160.

<sup>723</sup> Sophie Sparrow, *Describing the Ball: Improve Teaching by Using Rubrics – Explicit Grading Criteria*, 2004 MICH. ST. L. REV. 1, 6-15.

<sup>724</sup> BONE, *supra* note 681, at 11.

<sup>725</sup> See Menon, *supra* note 719, at 181.

the driving test. It is not important to establish whether more or less drivers pass this test in any one year (or at any one center) but only to ensure that the national pass standard is maintained.”<sup>726</sup> “[T]he implicit pedagogical philosophy underlying criterion-referenced assessment is that the fundamental purpose of professional education is not sorting, but producing as many individuals proficient in legal reasoning and competent practice as possible.”<sup>727</sup>

The use of clear criteria helps students understand what is expected of them as well as why they receive the grades they receive. Even more importantly, it increases the reliability of the teacher’s assessment by tethering the assessment to explicit criteria rather than the instructor’s gestalt sense of the correct answer or performance.<sup>728</sup> The criteria should be explained to students long before the students undergo an assessment. This enhances learning and encourages students to become reflective, empowered, self-regulated learners.<sup>729</sup>

#### 4. Use Assessments to Inform Students of Their Level of Professional Development.

**Principle:** The teacher uses assessments to inform students of their level of professional development.

**Comments:**

The development of expertise takes time, and there are stages with discernable differences: novice, advanced beginner, competent, proficient, and expert.<sup>730</sup> Therefore, our assessments should communicate to students where their development of professional expertise stands. Defining the level of proficiency that we want law students to achieve at each stage of their professional development is a task that warrants the attention of legal educators.

In communicating with students about their level of expertise in legal analysis, for example, one might want to articulate assessments for students in terms of levels of proficiency, perhaps linked to characteristics of student performance in the following way:

*Limited proficiency:* overly simplistic, incomplete analysis that misses key issues and fails to use relevant legal rules, facts and policy;

*Basic competence:* formalistic analysis that recognizes many issues, distinguishes relevant and irrelevant principles, and makes substantial but incomplete use of relevant rules, facts and policy;

*Intermediate competence:* integrated analysis that addresses nearly all issues, focusing on and developing relevant rules, facts and policy in a meaningful way that reflects conceptual understanding rather than a formulaic approach, and spots but does not work extensively or effectively with issues involving substantial uncertainty or novelty;

---

<sup>726</sup> BONE, *supra* note 681, at 4.

<sup>727</sup> SULLIVAN ET AL., *supra* note 7, at 210-11.

<sup>728</sup> Sparrow, *supra* note 723, at 28-29.

<sup>729</sup> *Id.* at 22-25.

<sup>730</sup> Wegner, Assessment, *supra* note 24, at 11.

*Advanced proficiency:* demonstrates characteristics of intermediate proficiency, but also considers implications of analysis more fully, brings to bear sound and creative approaches, works extensively and effectively with issues involving substantial uncertainty or novelty.<sup>731</sup>

Another way of indicating students' progress toward expertise is illustrated by the following scale that the Law Society of England and Wales requires Legal Practice Course providers to use. It not only indicates whether a student can perform a task, transaction, or skill, but also assesses the level of supervision that the student requires. "Course providers could then provide the student with a graduated record indicating the level of achievement demonstrated. The student should then be able to identify the level of supervision required in the future and be able to plan his or her future learning needs accordingly."<sup>732</sup>

- the student is familiar with the skill, task or transaction, but not able to perform it.
- the student can perform the skill, task or transaction, but requires closely supervised practice.
- the student can perform the skill, task or transaction with minimal supervision.
- the student can perform the skill, task or transaction adequately without further training.
- the student can perform the skill, task or transaction in an outstanding manner with virtually no supervision and could provide assistance to others.

Similar descriptions can be developed for any of the competencies that we want students to develop during law school. For example, an on-going project by faculty at Georgia State University College of Law, the Glasgow Graduate School of Law, and the Dundee Medical School is developing assessment criteria for evaluating lawyer-client communication skills, beginning with client interviewing.<sup>733</sup> The project breaks down the components of effective client interviewing skills into discrete segments with descriptions of various levels of proficiency.

The project's emerging assessment tool was used as part of the summative Interviewing Assessment at the Glasgow Graduate School of Law in January, 2006, which also involved standardized clients.<sup>734</sup> The analysis of data following that assessment indicated a close correlation among ratings of the interviews made by standardized clients, practicing lawyers serving as evaluators, and academic staff.<sup>735</sup>

The form used in the Glasgow assessment posed the following eight questions which were rated on a scale of 1 to 5.

1. The greeting and introduction by the student lawyer was appropriate.
2. I felt the student lawyer listened to me.
3. The student lawyer approach to questioning was helpful.
4. The student lawyer accurately summarized my situation.

---

<sup>731</sup> *Id.* at 12.

<sup>732</sup> Legal Practice Course, *supra* note 142, at 25-26.

<sup>733</sup> Karen Barton, Clark D. Cunningham, Gregory Todd Jones & Paul Maharg, *Valuing What Clients Think: Standardized Clients and the Assessment of Communicative Competence*, 13 CLINICAL L. REV. 1 (Fall 2006).

<sup>734</sup> *Id.* at 33-41.

<sup>735</sup> *Id.* at 41-50.

5. I understood what the student lawyer was saying.
6. I felt comfortable with the student lawyer.
7. I would feel confident with the student lawyer dealing with my situation.
8. If I had a new legal problem I would come back to this student lawyer.

Explicit criteria described how many points to award for each of the eight topics. For example, the following criteria were used for awarding points on number 6, “I felt comfortable with the student lawyer:”

- |           |  |
|-----------|--|
| 1 point:  | Lawyer was bored, uninterested, rude unpleasant, cold, or obviously insincere.                                   |
| 2 points: | Lawyer was mechanical, distracted, nervous, insincere, or used inappropriate remarks.                            |
| 3 points: | Lawyer was courteous to you and encouraged you to confide in him or her.   |
| 4 points: | Lawyer was generally attentive to and interested in you. You felt confident to confide in him/her.               |
| 5 points: | Lawyer showed a genuine and sincere interest in you. There was a sense of connection between you and the lawyer. |

Hopefully, more collaborations like the Glasgow/Georgia State project will lead to the development of additional descriptions of levels of proficiency in professional expertise and a growing consensus about what we should be teaching students and how we can measure success.

Our greatest challenge is finding effective ways to assess the overall competence of our students. If our program of instruction aims to develop competence, we should be concerned about how best to evaluate the level of competence of each student. In order to do this, we must put students in the roles of lawyers.

Legal analysis alone is only a partial foundation for developing professional competence and identity. It is not enough even to develop analytic knowledge plus merely skillful performance. The goal has to be integration into a whole greater than the sum of its parts. Assessment of students’ learning and growth need to be consistent with the goal of this integration: professional judgment and the ability to continue to learn and develop toward the highest standards of the legal profession. These broader aspects of professional development can be assessed in ways that can help students, but the assessment must take place “in role,” rather than in the more detached mode that “law of lawyering” courses typically foster.<sup>736</sup>

Assessments of competence would not only assess students’ knowledge and capabilities but also their professionalism. This is not easy to achieve, but the medical profession has demonstrated that it is possible.

Assessing the more complex goal of students’ professionalism or ability to embody good ethical and professional judgment is more

---

<sup>736</sup> SULLIVAN ET AL., *supra* note 7, at 225.

difficult to achieve . . . . Significant evidence from medical schools, however, suggests that some basic aspects of professionalism can be assessed and that, moreover, such assessments yield highly significant predictions about which students are likely to exhibit problematic behaviors as practitioners.<sup>737</sup>

Medical educators are much more advanced than legal educators in thinking about assessment issues and developing tools and methods to assess student learning. “Medical educators, hearing the call of public accountability, are adapting educational programs to teach apprentice practitioners in a way that ensures competent practice.”<sup>738</sup> We can learn from their experience.

The measurement of professional behavior is one of the greatest challenges in medical education today. Professional behaviors are very difficult to measure with paper-and-pencil tests because of the likelihood that students will respond with socially desirable, as opposed to personally realistic, choices. As a result, the best measures of professional behavior lie in the context of clinical activity and involve a conflict that the student or resident must resolve under supervision.<sup>739</sup>

According to Drs. Ronald Epstein and Edward Hundert,<sup>740</sup> the three most commonly used assessment methods in medical schools are subjective assessments by supervising clinicians, multiple-choice examinations, and standardized patient assessments.

1. *Subjective assessments by supervising clinicians.* During clinical experiences, faculty physicians observe students’ performance and rate them not only on their scientific and technical competence, but also on “dimensions of professionalism, including compassion, respect, interprofessional relationships, and conscientiousness.”<sup>741</sup> These ratings can lack reliability for numerous reasons.

[E]valuators often do not observe trainees directly. They often have different standards and are subject to halo effects and racial and sex bias. Because of interpatient variability and low interrater reliability, each trainee must be subject to multiple assessments for patterns to emerge. Standardized rating forms for direct observation of trainees and structured oral examination formats have been developed in response to this criticism.<sup>742</sup>

Another format being used to evaluate professional competence is to have trainees present several best-case videotapes of their performance in real clinical settings to a trained examiner who uses

---

<sup>737</sup> *Id.* at 222.

<sup>738</sup> David Stern, MD, PhD, *Outside the Classroom: Teaching and Evaluating Future Physicians*, 20 GA. ST. U. L. REV. 877, 903 (2004).

<sup>739</sup> *Id.* at 902 (citations omitted).

<sup>740</sup> Epstein & Hundert, *supra* note 150, at 226.

<sup>741</sup> Stern, *supra* note 738, at 902.

<sup>742</sup> Epstein & Hundert, *supra* note 150, at 230 (citations omitted).



specified criteria for evaluation.<sup>743</sup> “Although the face validity of such a measure is high and the format is well-accepted by physicians, the number of cases that should be presented to achieve adequate reliability is unclear.”<sup>744</sup>

2. *Multiple-choice examinations.* Multiple choice examinations have been proven to be a highly reliable way to evaluate factual knowledge and problem-solving skills and to assess some aspects of context and clinical reasoning.<sup>745</sup>
3. *Standardized patient assessments.* The use of standardized patients in an Objective Structured Clinical Examination (OSCE) can produce reliable ratings of communication, physical examination, counseling, and technical skills if there is a sufficiently large number of standardized patient cases and if criteria for competence are based on evidence.<sup>746</sup> “Although few cases are needed to assess straightforward skills, up to 27 cases may be necessary to assess interpersonal skills reliably in high stakes examinations.”<sup>747</sup> It is difficult to define pass/fail criteria for OSCEs, and there is a debate about whether to use standardized patients or external raters.<sup>748</sup> “The OSCE scores may not correlate with multiple-choice examinations and academic grades, suggesting that these tools measure different skills.”<sup>749</sup>

Peer ratings can provide accurate and reliable assessments of physician performance, especially professionalism.<sup>750</sup> Peers may be in the best position to evaluate professionalism; people often act differently when not under direct scrutiny. Anonymous medical student peer assessments of professionalism have raised awareness of professional behavior, fostered further reflection, helped students identify specific mutable behaviors, and been well-accepted by students. Students should be assessed by at least 8 of their classmates. The composite results should be edited to protect the confidentiality of the raters.<sup>751</sup>

Self-assessment is another tool that has helped evaluate the competency of physicians. Self-assessments have been used with some success in standardized patient exercises and in programs that offer explicit training in the use of self-assessment instruments. Among trainees who did not have such training, however, self-assessment was neither valid nor accurate. Rather, it was more closely linked to the trainee’s psychological sense of self-efficacy and self-confidence than to appropriate criteria, even among bright and motivated individuals.<sup>752</sup>

The various types of assessments make it difficult to rank students, because a student may excel in some dimensions and struggle in others. “However, one rarely

---

<sup>743</sup> *Id.*

<sup>744</sup> *Id.*

<sup>745</sup> *Id.* at 230 (citations omitted).

<sup>746</sup> *Id.*

<sup>747</sup> *Id.* (citations omitted).

<sup>748</sup> *Id.*

<sup>749</sup> *Id.* (citations omitted).

<sup>750</sup> *Id.* at 231.

<sup>751</sup> *Id.* (citations omitted).

<sup>752</sup> *Id.* (citations omitted).

needs this process of ranking in a field where competence, rather than comparative excellence, is the essential characteristic.”<sup>753</sup>

In the medical profession, many people are supporting the development of more comprehensive licensing examinations that add structured direct observations, OCSE standardized patient (SP) stations, real patient cases, case-based questions, peer assessments, and essay-type questions to the traditional computer-gradable formats.<sup>754</sup>

Comprehensive assessments link content across several formats. Post-encounter probes immediately after SP exercises using oral, essay, or multiple-choice questions test pathophysiology and clinical reasoning in context. Triple-jump exercises – consisting of a case presentation, an independent literature search, and then an oral or written postencounter examination – test the use and application of medical literature. Validated measures of reflective thinking have been developed that use patient vignettes followed by questions that require clinical judgment. These measures reflect students’ capacity to organize and link information; also, they predict clinical reasoning ability 2 years later. Combining formats appears to have added value with no loss in reliability.<sup>755</sup>

The website of the Accreditation Council for Graduate Medical Education (ACGME)<sup>756</sup> reflects an effort by that organization to assist medical professionals in expanding their repertoire of assessment tools and thereby expand the range and diversity of skills assessed. For example, the website details a range of tools for assessing students’ development of interpersonal and communication skills, including rating forms completed by patients, coding of videotaped patient interviews, and self-rating on a humanism scale.<sup>757</sup>

The ACGME Outcome Project’s TOOLBOX OF ASSESSMENT METHODS<sup>758</sup> includes descriptions and examples of instruments recommended for use by programs as they assess the outcomes of their educational efforts. These include:

1. *360-Degree Evaluation Instrument*. Ratings forms completed by supervisors, peers, subordinates, and patients and families to provide feedback about a person’s performance on several topics (e.g., teamwork, communication, management skills, decision-making).
2. *Chart-Stimulated Recall Oral Examination (CRS)*. A trained and experienced physician examiner questions the examinee about the care provided probing for reasons behind the work-up, diagnoses, interpretation of clinical findings, and treatment plans.

---

<sup>753</sup> Stern, *supra* note 738, at 903.

<sup>754</sup> Epstein & Hundert, *supra* note 150, at 232.

<sup>755</sup> *Id.* (citations omitted).

<sup>756</sup> ACGME Outcome Project, *supra* note 124, at Competencies to Assess, Complete List, <http://www.acgme.org/outcome/assess/complist.asp>.

<sup>757</sup> *Id.* at Interpersonal and Communication Skills Assessment Approaches, [http://www.acgme.org/outcome/assess/IandC\\_Index.asp](http://www.acgme.org/outcome/assess/IandC_Index.asp).

<sup>758</sup> *Id.* at Toolbox of Assessment Methods (Version 1.1 2000), <http://ACGME.org/Outcome/assess/Toolbox.pdf>.

3. *Checklist Evaluation of Live or Recorded Performance.* Checklists consist of essential or desired specific behaviors, activities, or steps that make up a more complex competency or competency component.
4. *Global Rating of Live or Recorded Performance.* Global rating forms are distinguished from other rating forms in that (a) the global rater judges general categories of ability (e.g., patient care skills, medical knowledge, interpersonal and communication skills) instead of specific skills, tasks, or behaviors; and (b) the ratings are completed retrospectively based on general impressions collected over a period of time (e.g., end of a clinical rotation) derived from multiple sources of information (e.g., direct observations or interactions; input from other faculty, residents, or patients; review of work products or written materials).
5. *Objective Structured Clinical Examination (OSCE).* One or more assessment tools are administered at 12 to 20 separate standardized patient encounter stations, each station lasting 10-15 minutes. Between stations candidates may complete patient notes or a brief written examination about the previous patient encounter. All candidates move from station to station in sequence on the same schedule. Standardized patients are the primary assessment tool used in OSCEs, but OSCEs have included other assessment tools such as data interpretation exercises using clinical cases, and clinical scenarios with mannequins to assess technical skills.
6. *Procedure, Operative, or Case Logs.* These logs document each patient encounter by medical conditions seen and surgical operations or procedures performed.
7. *Patient Surveys.* Surveys of patients to assess satisfaction with hospital, clinic, or office visits typically include questions about the physician's care. The questions often assess satisfaction with general aspects of the physician's care, (e.g., amount of time spent with the patient, overall quality of care, physician competency (skills and knowledge), courtesy, and interest or empathy). More specific aspects of care can be assessed including: the physician's explanations, listening skills and provision of information about examination findings, treatment steps, and drug side effects.
8. *Portfolios.* A collection of products prepared by the resident that provides evidence of learning and achievement related to a learning plan. A portfolio typically contains written documents but can include video- or audio-recordings, photographs, and other forms of information. Reflecting upon what has been learned is an important part of constructing a portfolio.
9. *Record Review.* Trained staff in an institution's medical records department or clinical department perform a review of patients' paper or electronic records.

10. *Simulations and Models.* Simulations used for assessment of clinical performance closely resemble reality and attempt to imitate but not duplicate real clinical problems. Key attributes of simulations are that: they incorporate a wide array of options resembling reality, allow examinees to reason through a clinical problem with little or no cueing, permit examinees to make life-threatening errors without hurting a real patient, provide instant feedback so examinees can correct a mistaken action, and rate examinees' performance on clinical problems that are difficult or impossible to evaluate effectively in other circumstances. Simulation formats have been developed as paper-and-pencil branching problems (patient management problems or PMPs), computerized versions of PMPs called clinical case simulations (CCX®), role-playing simulations (e.g., standardized patients (SPs), clinical team simulations), anatomical models or mannequins, and combinations of all three formats.
11. *Standardized Oral Examination.* A type of performance assessment using realistic patient cases with a trained physician examiner questioning the examinee. The examiner begins by presenting the examinee with a clinical problem in the form of a patient case scenario and asks the examinee to manage the case. Questions probe the reasoning for requesting clinical findings, interpretation of findings, and treatment plans.
12. *Standardized Patient Examination.* Standardized patients (SPs) are well persons trained to simulate a medical condition in a standardized way or actual patients who are trained to present their condition in a standardized way.
13. *Written Examination.* A written or computer-based MCQ examination is composed of multiple-choice questions (MCQ) selected to sample medical knowledge and understanding of a defined body of knowledge, not just factual or easily recalled information.

Other innovations that are being used to assess the professional competence of physicians include:

1. Multimethod assessment.
2. Clinical reasoning in situations that involve clinical uncertainty.
3. Standardized patient exercises linked to postencounter probes of pathophysiology and clinical reasoning.
4. Exercises to assess the use of medical literature.
5. Long-station standardized patient exercises.
6. Simulated continuity.
7. Teamwork exercises.
8. Unannounced standardized patients in clinical settings.
9. Assessments by patients.
10. Peer assessment of professionalism.
11. Portfolios of videotapes.
12. Mentored self-assessment.

13. Remediation based on a learning plan.<sup>759</sup>

It would be a worthwhile project for legal educators to investigate the feasibility of applying the techniques mentioned in this section to assessments during law school, as part of the bar examination, and after entry into practice.

5. **Be Sure Assessment is Feasible.**

**Principle:** **The teacher uses assessments to measure outcomes that are reasonably possible to assess validly, reliably, and fairly.**

**Comments:**

Feasibility is an additional consideration. There may be some desirable outcomes that are impossible or too difficult to assess. For example, it may not be feasible to assess a student's commitment to justice. This does not mean law schools should stop trying to instill a commitment to seek justice in students, but we may not be able to measure how well we are succeeding. Therefore, we should be careful to distinguish between desired outcomes and measurable outcomes.

On the other hand, if law teachers make the effort, we may discover ways to evaluate some things that we might initially consider unmeasurable. For example, Laurie Morin and Louise Howells believe they found a way to measure the development of students' reflective judgment.<sup>760</sup> We should closely monitor the progress of Marge Shultz and Sheldon Zedeck's effort to create a new Law School Admissions Test (see the section on various statements of desirable outcomes of legal education in Chapter Two). If they succeed in developing tests that measure some or all of the twenty-six factors related to effective lawyering, their project will have implications for assessing law student learning, not just their qualifications for law school admission.

There may be some desirable outcomes that we could assess, but it is not feasible to do so because of the time and training required to implement the assessment, equipment or technology required, number of assessments required per examinee, or financial cost. We should not stop trying to achieve desirable outcomes because they are difficult to assess, but we should be realistic about what we can assess and whether it is imperative that we do so.

6. **Use Multiple Methods of Assessing Student Learning.**

**Principle:** **The teachers use multiple methods of assessing student learning.**

**Comments:**

"A valid, reliable, and fair picture of the student's ability is much more likely to exist if the measures are done several times using different modes of evaluation."<sup>761</sup>

An assessment may take the form of a final exam, a test administered after

<sup>759</sup> Epstein & Hundert, *supra* note 150, at 232.

<sup>760</sup> Laurie Morin & Louise Howells, *The Reflective Judgment Project*, 9 CLINICAL L. REV. 623 (2003).

<sup>761</sup> Munro, *supra* note 700, at 238.

a unit of instruction is covered, a paper, an observation of performance, a discussion between student and teacher, portfolio (profile) reviews, or some other method of determining what a student has learned. Before selecting an assessment tool, we should be clear about the goals of the assessment and the purposes for which it will be used.

The problem-based essay exam is the primary assessment tool used by legal educators in the United States. New methods could improve the quality of our assessments of student learning. One of the reasons why law teachers do not conduct formative assessments or more frequent summative assessments is the length of time it takes to read and evaluate large numbers of problem-based essay exams. Therefore, improvements in law school assessment would be enhanced by finding alternative forms of assessing learning.

Greg Sergienko makes a persuasive case for expanding the use of multiple choice exams, including the results of his study demonstrating that multiple choice tests can be more sophisticated tools than essay questions for analyzing students' abilities to read facts and cases as well as their ability to apply an unfamiliar rule of law to a legal problem.<sup>762</sup> Sergienko and Wegner agree that even problem-based essay exams can be scored much more quickly if they are criteria-referenced.<sup>763</sup>

We should not, however, overlook the value of helping students develop self-assessment skills.

A most important aspect of assessment is student self-assessment. Throughout an attorney's professional life after law school, her success in practice will depend on the ability to self-assess professional performance, behavior, and attitudes. "An indispensable trait of the truly competent lawyer, at whatever stage of career development, is that of knowing the extent and limits of his competence: what he can do and what requires the assistance of others." [CRAMTON REPORT, *supra* note 280, at 8.] Yet law students are trained in a tradition in which all assessment is external so that she never must assess herself. Early in law school, students need to be taught the essentials of assessment and need to be introduced to self-assessment. They need to assess their own work and then compare their assessment with that of their instructor. They need feedback on their ability to self-assess so that they can improve. Teachers can provide students with assessment instruments that reflect explicit criteria for the performance so that the students can judge their own performance. As Cramton said, we should view legal education "in long-run terms as preparation for a lifetime career involving continuous growth and self-development over a forty-year period." [CRAMTON REPORT, *supra* note 280, at 10.]<sup>764</sup>

Students would benefit from instruction in and application of peer-assessment and self-assessment methods. Law schools should also explore expanding the involvement of teaching assistants in assessments, at least for helping

---

<sup>762</sup> Greg Sergienko, *New Modes of Assessment*, 38 SAN DIEGO L. REV. 463, 493-505 (2001).

<sup>763</sup> Wegner, *Assessment*, *supra* note 24, at 33.

<sup>764</sup> MUNRO, *supra* note 4, at 124.



provide feedback on formative assessments.

Computerized testing and scoring holds great promise for the future in providing formative and summative assessments. Existing technology can help prepare assessment tools and evaluate the results. For example, there is a web-based platform called “Cyber Workbooks” that allows faculty to publish their course materials by integrating learning outcomes such as critical thinking, applied reasoning, and creative problem-solving. The platform consists of an authoring tool for developing course modules with lessons, questions, and answers, a user website accessible by students with a user name and password, and an administrative site for generating reports and allowing faculty to evaluate course modules. The platform has built-in assessment features that will identify, measure, validate, and report on learning outcomes and identify student weaknesses, without any special training. The program will time, grade, and record student responses to minimize faculty time and burden.<sup>765</sup>

## 7. Distinguish Between Formative and Summative Assessments.

**Principle:** The teacher distinguishes between formative and summative assessments.

**Comments:**

It is important to know what we will do with the information our assessments will produce. The purpose of an assessment can be formative, summative, or both. Formative assessments are used to provide feedback to students and faculty. Their purpose is purely educational, and while they may be scored, they are not used to assign grades or rank students. A summative assessment is one that is used for assigning a grade or otherwise indicating a student’s level of achievement. “Summative assessment occurs at the end of a course of study and is primarily used for the purpose of making a final judgment of the student alongside his or her peers – final in the sense that (unless there are mitigating circumstances) it is how a student performs in this assessment that will be used to decide whether a student can proceed, e.g., to the next level of the course or be admitted to a vocational course.”<sup>766</sup>

## 8. Conduct Formative Assessments Throughout the Term.

**Principle:** The teachers conduct formative assessments throughout the term.

**Comments:**

As mentioned above, formative assessments are used to provide feedback to students and faculty. Their purpose is purely educational, and while they may be scored, they are not used to assign grades or rank students. Current practices in the United States are uneven and inadequate. Some law teachers give practice exams and others use a variety of techniques to find out whether students are learning what we think we are teaching. The norm, however, is to give a final exam at the end of the semester without conducting any formative assessments during the course.

---

<sup>765</sup> For more information about “Cyber Workbooks” go to <http://www.cyberworkbooks.com>.

<sup>766</sup> BONE, *supra* note 681, at 4.

Providing formative feedback to students ought to be the primary form of assessment in legal education.

Contemporary learning theory suggests that efficient application of educational effort is significantly enhanced by the use of formative assessment. For educational purposes, summative devices have their place primarily as devices to protect the public by ensuring basic levels of competence. Formative practices directed toward improved learning ought to be the primary forms of assessment.<sup>767</sup>

Formative assessments are especially important for first year students.

For many students what is needed is time – time to adjust, grapple with hidden difficulties, and gain an intellectual home – and assistance – feedback that lets them know where they stand and how to move ahead more quickly. But time and assistance are exactly what is missing. Instead, first-year students are ranked and sorted at the end of each semester with profound consequences for the rest of their lives.<sup>768</sup>

The authors of the Carnegie Foundation's report explained why formative assessment is critical for educating professionals.

[T]he essential goal of professional schools must be to form practitioners who are aware of what it takes to become competent in their chosen domain and equips them with the reflective capacity and motivation to pursue genuine expertise. They must become “metacognitive” about their own learning, to use the psychologists' term. This is why effective means of formative assessment are so critical for training professionals.<sup>769</sup>

Formative assessments also help teachers know whether their coverage of a topic is sufficient or whether they need to review the material again or present it in a different manner. Educational experts advocate assessing student learning throughout the learning process and afterwards for the purpose of determining how to improve instruction and whether to continue or discard it. “If it becomes apparent that all or most of the students fail to comprehend a particular area of a course or a particular point made by the professor, this data indicates that the problem may be attributable to the professor.”<sup>770</sup> This information allows us to make corrections before any failures to learn become real problems.<sup>771</sup>

Formative assessments can take many forms. Giving practice exams is one example. Assigning short homework problems that could be reviewed by teaching assistants is another. There are various forms of peer-assessment or self-assessment exercises that can be used in class or between classes. Self-scoring computer quizzes can be created to help students practice taking exams and evaluate their strengths

---

<sup>767</sup> SULLIVAN ET AL., *supra* note 7, at 242.

<sup>768</sup> Wegner, *Assessment*, *supra* note 24, at 31.

<sup>769</sup> SULLIVAN ET AL., *supra* note 7, at 217.

<sup>770</sup> HESS & FRIEDLAND, *supra* note 304, at 286.

<sup>771</sup> SMITH & RAGAN, *supra* note 197, at 338.

and weaknesses.

Various forms of classroom assessments are gaining popularity in law schools. “Classroom assessment’ focuses on ‘small scale assessments conducted continuously by . . . teachers to determine what students are learning in that class.’ Classroom assessment is integral to learning and valuable because it is so proximate in time, providing immediate feedback to teacher and student.”<sup>772</sup> After class, the teacher can quickly review the students’ responses, determine whether the students have learned the intended lessons, report the results to the students, and plan remediation if necessary.<sup>773</sup>

Barbara Glesner-Fines encourages law teachers to use classroom assessment techniques for improving student learning and helping students build self-regulated learning skills.<sup>774</sup> Glesner-Fines identifies traditional methods of classroom assessment, such as watching student non-verbal cues, polling students, pop quizzes, and “The Minute Paper.” Another technique is to have students during or at the end of class submit written answers to questions such as “What is the most important point you learned today?” “What was the muddiest point in . . . ?” Paraphrase the \_\_\_\_ [rule or holding].” “Give an example of \_\_\_\_.”<sup>775</sup>

Thomas A. Angelo and K. Patricia Cross’s seminal work on this subject describe fifty effective techniques to assess student learning and faculty teaching in the classroom, including the ones mentioned above.<sup>776</sup> According to Greg Munro, these include techniques for assessing prior knowledge, recall, and understanding. The following techniques can be employed successfully in virtually any class:<sup>777</sup>

1. *Misconception/preconception check:* This classroom assessment technique uncovers prior knowledge or beliefs that may hinder or block learning. For example, law students studying auto casualty insurance in an insurance class often believe that Uninsured Motorist coverage applies only when the insured is driving or riding as a passenger in a vehicle, when, in fact, the policy language covers the insured as a pedestrian hit by an uninsured motorist, which coverage accords with the legislative intent to protect the public from injury by uninsured motorists. Students also believe that Bodily Injury Liability coverage will provide benefits to a driver injured in a single-vehicle rollover, when, in fact, it covers the driver only for liability to others. These misconceptions can be revealed and dealt with by means of the misconception/preconception check before

---

<sup>772</sup> Munro, *supra* note 700, at 241 (quoting K. Patricia Cross, *Feedback in the Classroom: Making Assessment Matter* 5 (AAHE Assessment Forum, AM. ASSN. FOR HIGHER EDUC. 1988)). See also HESS & FRIEDLAND, *supra* note 304, at 261 (encouraging the use of classroom feedback as formative assessments).

<sup>773</sup> Barbara Glesner-Fines, *Classroom Assessment Techniques for Law School Teaching*, in ASSESSMENT, FEEDBACK, AND EVALUATION: EIGHTH ANNUAL CONFERENCE OF THE INSTITUTE FOR LAW SCHOOL TEACHING (2001). For information about using technology to facilitate asking questions and tabulating answers, see the section “Enhance Learning With Technology.”

<sup>774</sup> *Id.*

<sup>775</sup> *Id.*

<sup>776</sup> THOMAS A. ANGELO & K. PATRICIA CROSS, CLASSROOM ASSESSMENT TECHNIQUES: A HANDBOOK FOR COLLEGE TEACHERS (1993).

<sup>777</sup> The following descriptions of assessment techniques were copied almost verbatim from Munro, *supra* note 700, at 242-44 (citations omitted).

covering the material.

2. *Minute papers:* The “one-minute paper” or “half-sheet response” asks students in a couple of minutes or on a half sheet of paper some variation of the questions, “What is the most important thing you learned during this class?” and “What important question remains unanswered?” This allows the professor to assess whether students are getting the main theme around which the material is based or are meeting learning objectives. It also lets the professor know what students do not understand. This is especially important, since faculty often assume students have learned or have a base of knowledge when, in fact, they do not.
3. *Empty outlines:* The professor gives the students a partially completed outline and asks them to fill in the outline for the material covered in the reading, lecture, or other materials.
4. *Categorizing grids:* This technique requires students to sort information in appropriate conceptual categories.
5. *Defining features matrix:* This assessment matrix requires students to categorize concepts according to the presence or absence of certain defining features. For example, students in a securities or business regulation course might be asked to categorize transactions on a matrix defining features that determine whether the transaction constitutes a security for purposes of regulation. Students in an insurance class might categorize on a matrix various forms of contract to determine whether they are “insurance” for purposes of insurance regulation.
6. *Classroom opinion polls:* This device helps students to be aware of their own opinions, weigh them in light of those of their peers, and test them against evidence and expert opinion.
7. *Course-related self-confidence surveys:* The professor designs this survey with a few simple questions designed to determine the students’ self-confidence in an ability or skill. This allows the professor to evaluate the best approach to student learning and the needs of the students. For example, a professor in a trial advocacy class might design a survey asking students their level of confidence that, in this class, they will gain the ability to speak publicly, conduct voir dire, make a prepared statement of what their evidence will show, perform cross-examination, or make a closing argument. The survey may reveal that students lack confidence in their ability to cross-examine a witness or to carry on a voir dire dialogue with a jury. The professor can then work with students on strategies to overcome that lack of confidence.
8. *Electronic mail feedback:* The professor asks a single question by e-mail to the class. Each student responds with a personal, anonymous message to the professor’s electronic mailbox. This provides a fast method of receiving immediate feedback on an issue regarding

teaching or teacher.

9. *Group instructional feedback technique:* This method provides a peer reviewed but anonymous form of teaching evaluation.<sup>778</sup> Generally, a facilitator from outside the school visits the class, which has been divided into small groups. The facilitator asks the groups three questions regarding the course and instruction: (1) What works? (2) What does not work? (3) What can be done to improve the course or instruction? The facilitator then presides over reporting by the groups to help them arrive at consensus on the three questions. The facilitator reports the results to the professor, allowing the process to remain anonymous but providing valid, reliable, and fair feedback to the professor.

Angelo and Cross point out several positive characteristics of classroom assessment. They note that, although it is teacher directed, “depending on the judgment, wisdom and experience of the teacher,” it is simultaneously learner centered. Moreover, it is mutually beneficial to both teacher and students. Classroom assessment is formative, not designed to be “evidence for grading,” but part of the learning process. It is ongoing and can become part of the “daily feedback loop between students and teacher.”<sup>779</sup>

Technology is presenting some new ways to conduct classroom assessments. For example, classroom performance systems use “clickers,” in which each student is given a keypad to respond to in-class multiple choice questions. The software records and reports on the results as a tool for responding to students’ diverse ways of learning and serves as a classroom assessment technique that informs the teachers whether the students are learning and informs the students whether their learning strategies are working productively. Another technological innovation is the use of recording systems which automatically make video and sound records of students’ classroom answers and performances for subsequent review.

Legal educators should strive to provide students with formative feedback on their progress in every course before administering summative evaluations. Our students need it, and they deserve it.

### **9. Conduct Multiple Summative Assessments Throughout the Term, When Possible.**

**Principle:** The school conducts multiple summative assessments of student learning throughout the term, when possible.

**Comments:**

Although law school exams provide a mechanism for assigning grades and ranking students, a single examination is an inadequate tool for determining which students have learned and which have not. The stakes of evaluation are high: grades serve to rank students for prospective employers and reflect on students’

---

<sup>778</sup> For a more detailed description, see Gregory S. Munro, *More Effective Evaluation of the Course and Instructor*, in HESS & FRIEDLAND, *supra* note 304, at 281.

<sup>779</sup> Munro, *supra* note 700, at 244 (citations omitted).

chances for admission to other educational programs.<sup>780</sup> Multiple evaluations of student learning increase the accuracy of the conclusions about student performance, improve student performance on the final examination, and increase the range of skills, values, and knowledge that the instructor may evaluate.<sup>781</sup>

A single assessment has significant potential for error because a student might be ill or have other personal issues that can distort the accuracy of the evaluation. The potential for distortion is exacerbated by the fact that a single assessment produces higher levels of stress because of its significance to the student's grade in the course and future. Similarly, there is a greater potential for teacher error if only one summative assessment is administered per term, particularly when problem-based essay exams are used.

There may be some justification for delaying summative assessments to the end of the semester if it would be unfair to evaluate students earlier. For example, first year students' analytical skills may not be sufficiently developed until the end of the first semester, or even the first year, to administer summative assessments sooner. However, it may be that some aspects of first year learning should be summatively assessed during the term, particularly students' understand of legal doctrine or their ability to read and understand appellate cases (both of which could be assessed with multiple choice tests). In upper level courses where the transmission of legal doctrine or other knowledge is a significant objective of the course, there is no excuse for not conducting summative assessments throughout the term.

Legal educators in the United States should also reconsider the current practice of allowing individual professors to draft and grade their own exams without any oversight. Summative assessments should be collaboratively created and graded, as is the common practice in British Commonwealth jurisdictions. To the extent that resources permit, summative assessments should be vetted by learning experts, at least periodically.

#### **10. Ensure That Summative Assessments Are Also Formative Assessments.**

**Principle:** The school ensures that summative assessments are also formative assessments.

**Comments:**

Students cannot learn unless the results of their summative assessments are explained to them. Assigning a student a grade or even describing the level of professional development does not help the student understand how to improve. For example, a summative evaluation might indicate that a student's performance on an exam demonstrated *limited proficiency*, that is, it showed overly simplistic, incomplete analysis that misses key issues and fails to use relevant legal rules, facts and policy. This conclusion, however, does not provide any basis upon which the student can understand the shortcomings of the student's analysis or how it could be improved. As the ABA's Task Force on Lawyer Competency recommended in 1979, "[l]aw schools and law teachers should develop and use more comprehensive methods of measuring law student performance than the typical end-of-the-term examination.

---

<sup>780</sup> HESS & FRIEDLAND, *supra* note 304, at 285.

<sup>781</sup> *Id.* at 290.



Students should be given detailed critiques of their performance.<sup>782</sup> Students learn with feedback.

In American law schools, final exams are not returned to students unless students ask to see theirs, and most law teachers do not try to explain the results of final examinations to the entire class. This tradition is inconsistent with best practices, because it misses an opportunity to use final examinations to enhance student learning.

How did the student answer the question? Did he grasp the problem? Did he analyse the facts properly? Did he argue effectively? What are his weaknesses? The student never knows. I have in my files 143 [scripts]; not one of my students knows anything other than the final mark. The [scripts] were not to be returned. I defy anyone to tell me this is proper educational process . . . . [M]arginal comments on the returned [script] would certainly serve as a teaching device. Individual or group discussion of the examination should be part of the teaching process.<sup>783</sup>

Teachers should return all written exams and papers to students, with notes indicating specific strengths and shortcomings. Teachers should explain to students how they fared on other forms of summative evaluations. Teachers should provide model answers to exams, and encourage students to seek guidance about how to improve, either through internet correspondence, one on one meetings, class debriefings, or other methods.<sup>784</sup>

Michael Hunter Schwartz developed a form designed to facilitate student reflection and self-regulation with respect to law school exams, papers and other graded work. The form asks the student to: (1) compare how well she did with how well she expected to have done (to improve student self-assessment), (2) identify what she did incorrectly, in part by identifying the professorial comments most frequently appearing on her paper, (3) identify the causes of any errors in her work, focusing on correctable causes such as incorrect learning strategy choices or insufficient persistence, and (4) plan how she will avoid the error(s) in the future.<sup>785</sup>

## 11. Require Students to Compile Educational Portfolios

**Principle:** The school requires students to compile educational portfolios.

**Comments:**

Educational portfolios are seldom used in the United States, although they are required throughout the system of higher education in the United Kingdom. They can take many forms, but essentially they are compilations of materials that

<sup>782</sup> CRAMTON REPORT, *supra* note 275, at 4.

<sup>783</sup> BONE, *supra* note 681, at 15 (citing Albert Orschel, *Is Legal Education Doing its Job?*, 40 ABA J. 121, 124 (Feb. 1954)).

<sup>784</sup> See Richard Henry Seamon, *Lightening and Enlightening Exam Conferences*, 56 J. LEGAL EDUC. 122 (2006) (describing how exam conferences can help students learn the law, write better exam answers, and avoid discouragement and cynicism and how they can help faculty teach better, write better exam questions and grade them more fairly and accurately, and avoid discouragement and cynicism).

<sup>785</sup> Exercise 16-2, SCHWARTZ, *supra* note 406.

document a student's academic achievement and personal development.<sup>786</sup> Their perceived benefits include "making the results of learning in higher education more explicit, placing greater responsibility on students to understand and direct their own learning and personal growth, integrating academic and extracurricular development, creating more effective means to track student progress and enhance program quality, and assisting students in their search for employment."<sup>787</sup>

A system of student portfolios addresses many of the assessment principles previously discussed. They focus both teachers' and learners' attention on learning, and make multi-faceted learning that progresses throughout students' educational lifetimes visible in fresh and meaningful ways. Portfolios place responsibility squarely on learners to consider how diverse academic and outside learning relate, and bring them into closer, meaningful contact with advisers who can monitor and encourage their work. Portfolios also provide a convenient means both for documenting professional capability in the interest of future employment, and encouraging and recognizing distinguished work.<sup>788</sup>

Portfolios can be particularly helpful for students who do not get off to the best start but whose expertise and academic achievement mature as they proceed through law school. In fact, portfolios can facilitate a student's development by causing the student to reflect on her personal and professional objectives and by providing a tool for demonstrating that first semester grades do not accurately reflect her potential as a lawyer. Examples of materials that might be included in a student's educational portfolio include:

short reflective essays on personal and professional goals at the start of law school and each successive year; writing samples and work product of various sorts; resumes; certificates of academic distinction awarded for advanced proficiency in the first year or honors performance thereafter; "learning logs" associated with certain courses or work experiences; evidence of extracurricular activities that demonstrate effective work in teams or special professional contributions; statements regarding volunteer service of various sorts; letters of reference from faculty or work supervisors; evidence of research skills and use of advanced technology; and transcripts.<sup>789</sup>

A student might also use the portfolio to demonstrate her progress toward

---

<sup>786</sup> AMERICAN ASSOCIATION FOR HIGHER EDUCATION, *ELECTRONIC PORTFOLIOS: EMERGING PRACTICES IN STUDENT, FACULTY, AND INSTITUTIONAL LEARNING* (2001). While he was at Western State University College of Law, Michael Hunter Schwartz, now at Washburn University School of Law, designed and led a portfolio assessment process. The faculty identified the skills, knowledge, and values that Western States students should possess upon graduation, created a curriculum map identifying where in the curriculum students are introduced to, practice and must master those skills, knowledge, and values, and required students to create electronic web portfolios to which they submit evidence of attainment of the skills, knowledge, and values and reflect on those submissions. Schwartz expects to complete a law review article dealing with this project by the fall of 2006.

<sup>787</sup> Wegner, *Assessment*, *supra* note 24, at 70. See also Aaronson, *supra* note 33, at 6-7 (discussing the content and benefits of student portfolios).

<sup>788</sup> Wegner, *Assessment*, *supra* note 24, at 72-73.

<sup>789</sup> *Id.* at 71.

developing the fundamental skills and values needed for law practice. If a school records student performances in simulation-based courses or competitions, copies of a student's best performances could be included in the portfolio. This is made all the easier if the school uses digital recording devices. In fact, in this digital age, the entire portfolio can be electronic.<sup>790</sup>

Students would be able to provide selected materials from the portfolio to prospective employers, and schools could consider giving special academic recognition to students whose portfolios demonstrate outstanding achievement. "Criteria for such recognition would be made available well in advance to all interested students. Criteria would ideally be developed by faculty, in consultation with students, using the opportunity to articulate the meaning of excellence in light of the school's mission and goals, and the aspirations and potential of its students."<sup>791</sup>

---

<sup>790</sup> See BARBARA L. CAMBRIDGE, *ELECTRONIC PORTFOLIOS: EMERGING PRACTICES IN STUDENT, FACULTY, AND INSTITUTIONAL LEARNING* (2001).

<sup>791</sup> Wegner, *Assessment*, *supra* note 24, at 72.

