

## RESPONDING TO “EDUCATING LAWYERS”: AN HERETICAL ESSAY IN SUPPORT OF ABOLISHING TEACHING EVALUATIONS

*Dennis R. Honabach\**

THE quality of legal education has been under attack for some time and numerous fixes have been suggested. Nevertheless, if our goal is to adequately prepare our students for the practice of law, the Carnegie Foundation’s recently released report, *Educating Lawyers: Preparation for the Profession of Law*,<sup>1</sup> suggests we still have a long way to go. The intriguing question is why, after all this time, we law professors, as members of a profession that takes great pride in being problem solvers, have seemingly failed so miserably at improving legal education.

Much of the criticism of legal education has focused on our inability to teach well. Quality teaching, we are assured, will lead to quality legal education. We are told we should focus our efforts on being good teachers. We are urged to step up our efforts to evaluate faculty teaching abilities. Focusing more on teaching might solve the problem, but I think not. Indeed, I think the problem lies in the fact that we focus too much, not too little, on whether we are good teachers. I believe we care too much about teaching evaluations. Focusing on these evaluations undercuts our efforts to improve legal education. So strong is my belief that I propose the unthinkable. I propose we abolish teaching evaluations entirely!

The use of teaching evaluations has long been a matter of debate within the academy. While the precise method for evaluating teaching varies widely from one school to another, the general pattern involves both student assessment and peer assessment. Regardless of the form teaching evaluations take, many critics devalue them, question their reliability, their validity, or both.<sup>2</sup> While many of the perceived faults in the assessment process are real and might alone support a

---

\* Dean and Professor of Law, Chase College of Law, Northern Kentucky University. I thank Rick Bales, Nancy Firak, and Michael Hunter Schwartz for reading and making numerous helpful comments on earlier drafts of this essay.

1. WILLIAM M. SULLIVAN ET AL., *EDUCATING LAWYERS: PREPARATION FOR THE PROFESSION OF LAW* 185-202 (2007) [hereinafter *Carnegie Report*]. A summary of the Carnegie Report is available at [http://www.carnegiefoundation.org/files/elibrary/EducatingLawyers\\_summary.pdf](http://www.carnegiefoundation.org/files/elibrary/EducatingLawyers_summary.pdf) (last visited Oct. 14, 2007). The recently released book, *Best Practices for Legal Education*, makes many of the same criticisms. ROY STUCKEY ET AL., *BEST PRACTICES FOR LEGAL EDUCATION* 7 (2007).

2. See, e.g., Deborah Jones Merritt, *Bias, the Brain and Student Evaluations of Teaching*, 81 ST. JOHN’S L. REV. (forthcoming 2008), available at [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=963196](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=963196).

call for the abolition of teaching evaluation, those criticisms miss the crucial point entirely.

To understand why, we must first consider the general process of evaluating teaching. At most law schools, student assessments are administered using a predictable procedure. Sometime near the end of the semester, faculty members distribute an evaluation form containing a number of statements ranging from those asking the mundane (e.g., “The faculty member distributed a syllabus.”), to those asking the students to assess the content of the course (e.g., “The faculty member challenged me to think about policy issues relevant to the course.”), to the ultimate judgmental questions (e.g., “The faculty member is knowledgeable in the field.” and “I would recommend this faculty member or this course to others.”). Students indicate their agreement with the statement as it applies to their professors on a numeric scale, often running from “1” to some number, with higher numbers generally associated with “good” teaching and low numbers indicating sub-par performance.<sup>3</sup>

Somewhere on the form, students are offered the opportunity to include written comments explaining their evaluations. Often the students are prompted by specific questions such as, “Did the faculty member demonstrate respect for the students?” or “Was the text adequate?” Students are almost always offered the opportunity to make additional comments. Precious few, however, actually make any comments and those who often restrict their views to pithy statements such as “Professor So & So rocks” (or the opposite). The law school then sends those forms to be tabulated. The dean receives a printout for each course indicating the numerical average of the students’ responses to each of these questions. Sometimes even these averages are then averaged, producing some magical overall number. Individual faculty members are then given the opportunity to review their ratings after submitting their final grades for the course. The members of the faculty committee charged with evaluating untenured faculty also review the forms. Sometimes a summary of some sort is shared with all members of the faculty.<sup>4</sup>

Criticisms of these student evaluations, or “preference evaluations” as they are sometimes called, are legion. Some critics argue that students are not actually interested in undertaking a serious assessment of a course as evidenced by the fact that students can be enticed to do so only if the form is kept short, and the evaluation is done during class time. Second, under the standard process, students evaluate a course before it is over. Even those students who seriously attempt to assess a course must do so prior to preparing for the final, a process that some professors consider to be a crucial part of the learning process. Some critics question how students can accurately evaluate how much they have

---

3. At times the rating pattern is reversed, rather like scoring in golf, so that a score of “1” is excellent, a higher score being less desirable. The importance of identifying in which direction the scale runs was brought home to me painfully some years ago when I moved from a “high score is good” school to a “low score is good” school. Had I checked before hand, I would not have been so devastated by average scores of 1.5 or 2!

4. Except in some public schools, faculty members rarely get to see the student evaluations of their colleagues.

learned in the course without taking the final, an act that will give them perhaps their first real insight into how much they have learned. We ask students to evaluate the course before we allow them to assess the congruence between the final examination and the course or the quality of feedback they receive from their professors about their performance on the examination.<sup>5</sup> And, as critics are quick to point out, we ask students who often know little about a subject, except what they studied in courses, to assess a professor's knowledge of the field.<sup>6</sup> These shortcomings, critics maintain, raise questions about the validity of the evaluations.<sup>7</sup>

Even more critical are claims that students use the evaluation process to punish demanding, low-grading professors and, at the same time, to reward undemanding, high-grading professors. Consequently, to survive the tenure gauntlet and to attract students to their electives, professors enter into a tacit agreement with students: "I will dumb down my class and not ask much of you. You will take my course and not give me bad student evaluations."

Finally, there are concerns that student assessments reflect biases students might bring to the classroom. Do male teachers tend to receive higher grades than their female counterparts who are equally or more talented? Do professors of color systematically fare poorer on assessments? Do other factors such as age, weight, height, or style of dress affect assessments? Does faculty rank or administrative position play a role?

In sum, critics question the validity of student evaluations on three grounds. The students lack the incentive or the knowledge necessary to make a valid assessment of the faculty member. Evaluations reflect student biases. Finally, students use the evaluation process to make life easier for themselves, even at the expense of their education.

Are these criticisms valid? I am not convinced by the claim that students actually use the evaluation process to punish excellent, demanding, and low-grading teachers and to reward undemanding, easy-grading teachers. Indeed, I suspect that students are as likely to score a talented, demanding, low-grading professor as highly as they grade a talented, but less demanding higher-grading professor. I believe, however, that students may score ineffective, demanding, low-grading professors much more harshly than ineffective, undemanding, but high-grading professors. On the whole though, I suspect that students prefer effective teachers over ineffective teachers regardless of teacher demands and grading styles.

Some of the other criticisms of student evaluations may have more validity. Students probably lack a sound basis for knowing how much or little a professor

---

5. I thank Professor Michael Schwartz for this observation.

6. It is obvious to every law professor that no student can accurately assess the quality of thought in any particular field, except of course those few law review students at another law school who have accepted that professor's most recent article for publication in their law school's journal. Fortunately those students, however, do not distort the general picture described above because they, after all, evaluate someone else's teaching or something like that.

7. That criticism rings most loudly in those instances when we discover that the students cannot even agree if the professor distributed a syllabus or arrived to class on time!

knows. We certainly should only ask students about matters that they are capable of assessing. Moreover, we should recognize that using a single form for all courses is unlikely to allow students to validly critique a course. The use of a one-form-fits-all assessment tool flies in the face of our frequent statements that different faculty members have different pedagogical goals in different courses.

The question of whether student biases affect faculty evaluations is unanswered, although Professor Merritt makes a strong case in support of the claim.<sup>8</sup> My intuition tells me that student biases play a role in how students rate their professors, but my intuition has misled me many times.

Peer assessment, the second component of most teaching evaluations, may be even less reliable or valid. A standard criticism of peer assessment is that the faculty members who are charged with doing the assessments have no clear (or at least no clearly articulated) standards for evaluating teaching. No one, the canard goes, can define what good teaching is, and so evaluators have no idea of what is that they are looking for.<sup>9</sup> Indeed, the belief that the definition of *good teaching* is intractable has led some faculty members to conclude that we ought to not even attempt to assess the teaching abilities of our colleagues.

Even those who believe peer assessments may be valuable question the value of the evaluations that actually occur. They note several potential flaws. Classroom visits are terribly short, often lasting less than the length of even a single class period.<sup>10</sup> Even an entire classroom period is an awfully short time to assess the quality of any professional's work. The classroom experience is not constant from day to day. Indeed, different things happen at different times. Some faculty members teach in a one-period cycle and others employ a multi-day or even multi-week cycle. Whatever the cycle, making a judgment about a colleague's teaching ability after seeing only a short segment of a class makes little sense.

Moreover, the classroom experience is only a small fraction of what constitutes a course. The syllabus, classroom materials, texts, and other handouts that a faculty member chooses or provides are crucial parts of the course experience. Tests, projects, papers, and other evaluative materials the professor employs are equally important. So is the feedback on the student's performance. We all recognize that these factors greatly influence the effectiveness of our teaching. That is why so many faculties offer in-service programs on how to use technology and create evaluative tools. Yet, at how many tenured faculty meetings where we discuss a faculty member's teaching ability, do we hear a

---

8. See Merritt, *supra* note 2.

9. I have heard this criticism time and time again, particularly from untenured faculty members concerned about the uncertainty of such assessments. I should add that almost all of those very same individuals—at least those who survived the tenure process—feel much more confident later in their careers about their ability to assess the quality of the teaching of junior faculty members.

10. Admittedly, these situations tend to have involved favorable assessments of a colleague's teaching, suggesting that the assessment was actually nothing more than an opportunity to confirm a preexisting belief that the faculty member in question was a "good teacher." The unwritten rule seemed to be that to voice an unfavorable review, the evaluator needed to sit through almost an entire class period!

presentation on the quality of the faculty member's examination, teaching materials (e.g., web pages, handouts, etc.), or the feedback provided to students? At darn few, I bet, if my experiences are at all representative.

Almost any combination of these objections to the validity of teaching evaluations, student or peer, might well support abolishing their use. My reason for calling for their abolition, however, is quite different. I would abolish teaching evaluations because, simply put, we should not care whether we are *good* teachers. We should care whether our students are learning what we want them to learn. Our focus should be on our students and not on us. Whether we are good or bad teachers is the wrong question. *The right questions are whether we are effective teachers, and whether our participation in our students' education adds sufficient value to their learning experience to justify the tuition we charge.*<sup>11</sup>

The problem with focusing on *teaching evaluations* is that doing so distracts our attention from those crucial questions. As their very name indicates, the label "teaching evaluations" causes us to focus our attention on the teacher. The label urges us to pay attention to the teacher's use of various teaching techniques, and, of late, to the teacher's use of technology such as PowerPoint, the Internet, and clicker systems.<sup>12</sup> In short, the label "teaching evaluations" invites us to evaluate our performance as teachers. As a result, "good" teaching becomes an abstract concept independent of the effect that teaching has on the students in the class. When combined with the disagreements about the validity and reliability of current teaching evaluations, the tendency to view good teaching in abstract terms fails to drill down the important question of just how effective we are as teachers.

One might ask a number of questions about effective teaching. So what? Isn't a stimulating lecture a good lecture, even if the students did not understand the vocabulary? Isn't an extraordinarily vibrant Socratic give-and-take between a law professor and a few students an example of good teaching regardless of what the other students are doing in the class? Are not the lecturer and discussion leader filling their professional obligations? Don't we want to encourage such teachers to give those lectures and lead those discussions? The answer, of course, is "NO," at least if the students in those classrooms are not learning. That point, however, can easily become lost in a discussion about good teaching. Far too often, we debate the merits of the teaching performance while doing nothing to assess the effectiveness of the lecture or the discussion.

Even worse, when we attempt to undertake any assessment of the students' learning<sup>13</sup> and find the results wanting, we often adopt the age old tactic of blaming the victim, in this case the student. Our excuses are numerous. The class failed because: the students did not prepare adequately for the class; they

---

11. I add this last point because I am confident that most of our students are quite capable of learning a great deal without our aid. If we only asked ourselves whether they were learning anything, we could applaud ourselves for what they can achieve without our intervention.

12. In that sense, teaching evaluations are yet another traditional type of the input-based assessments that seem to bedevil legal education.

13. Here I am thinking of the student's performance on the final examination.

find the subject to be a particularly difficult one; they were ill-prepared for legal education; they simply do not care about learning; or whatever!<sup>14</sup> So long as we keep our focus on the teacher, we can deflect criticism and declare a class session an example of excellent teaching whether or not the students learned anything.<sup>15</sup> The students' failure to demonstrate that they learned what they should have can be chalked off as just that—the students' failure, not the faculty member's failure.

Focusing on the quality of teaching is just another example of the misguided belief that one can improve the quality of law school education by emphasizing educational inputs. The lack of any meaningful correlation between inputs and outputs, however, should seem self-evident.<sup>16</sup> One would have expected law faculties to have realized that long ago and to have changed their ways. Apparently, we have not.

Why have we paid more attention to faculty teaching than student learning? The reasons are numerous. In a semester-long course, we expect students to acquire a great deal of information. Creating a comprehensive examination that fairly tests the extent to which our students have mastered that information is itself somewhat problematic. Moreover, because information transmission is rarely the single goal in any of our courses, we would need to create an assessment tool that allows our students to demonstrate the degree to which they have mastered the skills we expect them to have acquired or sharpened in our courses.<sup>17</sup> And, if we really believe what we so often espouse, we would also need to formulate questions that would allow us to tease out the degree to which we have successfully imparted the values we wanted our students to internalize. I am not sure how we would accomplish all those tasks in a single examination, but I feel pretty confident that a few typical fact-situation essay questions will not do the trick. Adding a bank of multiple choice questions will only slightly improve the situation.

More problematic for our attempts to assess the effectiveness of our teaching is the simple fact that before we can adequately assess our effectiveness as teachers, we need to identify our teaching goals.<sup>18</sup> Our colleagues elsewhere

---

14. The irony with all of these "explanations" is that we who have been teaching for long periods of time made these same excuses to explain disappointing student performances when the crop of current new law teachers were the students about whom we were despairing!

15. Indeed, in most instances that is exactly how we do peer assessment. If you recognize that all of this sounds like the old saw that "the operation was successful but the patient died," you understand.

16. Not many would argue, for example, that the quality of legal education is directly related to the size of a law school's library, a law school's building, or even the size of a law school's faculty. See Theodore P. Seto, *Understanding the U.S. News Law School Rankings*, 60 SMU L. REV. 493, 558 (2007).

17. A classic example might be the common statement by faculty members teaching a first-year course who assert that they are teaching students how to read a case, but then never assess their students' ability to read a case. Students quickly learn through the grapevine that developing the ability to read cases is not tested and not surprisingly, they consequently spend little time mastering that skill.

18. That is, we can't do so unless we are willing to adopt the approach the courts have taken towards pornography—"we 'know [effective learning] when we see it.'" *Paris Adult Theatre I v.*

in the academy are moving, however slowly, toward adopting outcome-based education.<sup>19</sup> But, I dare say, few law professors begin a course by distributing a syllabus that details the learning objectives of their course. Some syllabi still amount to little more than a list of assignments. Those that state learning objectives generally do so in quite general terms such as, “in this course you will learn the basic rules of tort law” or “you will learn to apply the basic doctrines of corporate law to complex factual situations.” Those types of statements provide little guidance when we begin drafting our examination to assess our effectiveness (or to the students as they go about their work in the course). Without such guidance, it is difficult for us to discern whether we have achieved congruence between our learning goals, instructional methods, and methods of student assessment, the hallmark of effective education.

A third barrier to outcome assessment is the culture of the legal academy itself. Formulating learning objectives and assessing learning outcomes is difficult, and time consuming work. Even though most law school professors want to be good, indeed most want to be excellent, teachers, they are participants in an academic endeavor that values scholarship much more highly than it does teaching. Indeed, differences in merit raises tend to reflect scholarly productivity more than effective teaching. For example, summer stipends are more readily available to promote new scholarship than to improve teaching.<sup>20</sup> Finally, the production of cutting edge scholarship is the key to a faculty member’s mobility. Good scholars are in demand, effective teachers are not. Law schools will often hire a well-known scholar away from another school in an attempt to gain prestige, impress alumni, and raise the “intellectual atmosphere” of the law school. They almost never hire away a faculty member from another school because that faculty member enjoys a reputation as an excellent teacher.<sup>21</sup> Why would they, if no one can challenge their assertions that all their faculty members are good teachers? In sum, virtually all of the external incentives available to faculty members encourage faculty to spend their time producing scholarship.

Unfortunately, the internal incentives drive many law professors to also emphasize producing scholarship. New professors want to be effective teachers.

---

Slaton, 413 U.S. 49, 84 (1973) (Brennan, J., dissenting) (quoting *Jacobellis v. Ohio*, 378 U.S. 184, 197 (1964)).

19. Regional accrediting agencies appear to be among the agents driving this development. See, e.g., Jacob Ludes III, Executive Dir./CEO, New England Assoc. of Sch. & Colls., Inc., *The Role of Accreditation in Assuring Quality & Accountability in the Nation’s Schools & Universities*, Excerpts of Briefing to Legislative Staff of the Health, Education, Labor & Pensions Committee U.S. Senate (Jan. 14, 2003) (excerpts of briefing available at [http://www.neasc.org/neasc/briefing\\_senate.pdf](http://www.neasc.org/neasc/briefing_senate.pdf)).

20. Even when summer stipends are available for “teaching development,” their availability is often conditioned on the production of a tangible product such as an article on pedagogy. See Lee B. Vaughn, *Integrating Alternative Dispute Resolution (ADR) into the Curriculum at the University of Washington School of Law: A Report and Reflections*, 50 FLA. L. REV. 679, 694 n.40 (1998).

21. Admittedly, a top law school may refrain from extending an offer to a top law school professor who receives consistently poor teaching evaluations from her students but—as discussed above—student evaluations may be poor measures of teaching effectiveness. At the same time some of the start-up law schools such as Elon University School of Law appeared to have placed heavy weight on the teaching abilities of the seasoned faculty they recruited.

They hope their students will perform well on examinations. Most new professors, however, are discouraged by their students' performance on the first examinations they administer.<sup>22</sup> At this point, senior faculty members could play an important role in helping new faculty members understand how to improve their students' performance. Senior faculty members could assure new teachers that student performances will improve if the new faculty members work hard at becoming effective teachers. Senior faculty members could help new faculty members improve their courses.

Sadly, such mentoring is rarely the case. New faculty members are much more likely to be told that student performance on examinations is not an accurate assessment of student learning or a reflection of the professors' prowess. "Get over it," they are likely to be told. "Your expectations are too high. You can't expect that much from our students."<sup>23</sup> While many seasoned faculty members will volunteer to help young faculty members improve teaching skills, they rarely focus on student learning assessments.<sup>24</sup> Indeed, rather than being encouraged to investigate why their students performed poorly, some new teachers are actually told to make sure they create easy-to-grade examinations. That way, they can finish grading quickly and move on to more important things, i.e. scholarship! New faculty members soon discover that being named "Professor of the Year" is not very prestigious and that those who do are likely to be derided by their colleagues for "dumbing down" their course to garner student adulation. They learn that rather than being applauded for success, a faculty member whose students all earn high grades should expect to be accused of pandering to students. As a result, new faculty members find themselves adrift without any standards against which to measure their performance as educators. It does not take long for new faculty members to conclude that the safe route is to become a "Professor Goldilocks" whose grades and student evaluations are neither too good, nor too bad.

The dearth of external incentives for developing excellence as a teacher and the absence of internal incentives caused by the lack of benchmarks against which to measure one's effectiveness as a teacher have predictable effects on new professors. New teachers quickly learn that there is little to gain from developing valid and reliable outcome assessments of their students' learning. Improving their teaching to maximize their students' performance hardly seems worth the effort. On the other hand, they realize they have much to gain from being scholars. Thus, they publish and publish.<sup>25</sup>

---

22. See, e.g., Jay Feinman & Marc Feldman, *Pedagogy and Politics*, 73 GEO. L. J. 875, 881 (1985).

23. See, e.g., Michael Hunter Schwartz, *Teaching Law Students to Be Self-Regulated Learners*, 2003 MICH. ST. DCL L. REV. 447, 449 (describing that the primary reason students fail to perform better on law school examinations is the students' lack of ability).

24. New teachers are able to avoid the discomfort that comes with assigning low grades because most law schools have adopted some form of grade norming. Student performance in the course has no effect on the number of high grades to be given and a faculty member who gives all high grades should expect to be accused of pandering to students.

25. Or slink off to unknown areas where they sulk about the disproportionate weight given to scholarship.



So strong has the culture of scholarship become in legal education that one can rarely attend a discussion on scholarship these days without hearing someone espouse the belief that scholarship is essential for good teaching. Indeed, some observers have advanced the belief that the best scholars are actually the best teachers.<sup>26</sup> I know of no study actually demonstrating a positive relationship between scholarship and teaching. It seems unlikely that good scholarship and good teaching automatically go hand in hand.<sup>27</sup> What great scholars write about is rarely the stuff that makes for good learning in the classroom. If scholars really taught what they write about, most students—being novice learners—would simply be lost. While many great scholars are great teachers, at least some are not. Similarly, I suspect that some professors who rarely publish are quite effective teachers. Great scholarship and effective teaching are simply independent variables.<sup>28</sup> The emphasis on scholarship that is prevalent in most law schools may be useful for advancing our knowledge of law, but its impact on the effectiveness of law school teaching is likely to be neutral at best.

A second fact at play in the law school culture that undercuts the promotion of effective teaching is the general bias that many faculty members share against grading “soft skills.” Many faculty members eschew assigning grades to student skill performances. In many schools, anyone having the temerity to suggest that performance courses be graded just like any traditional course is likely to be confronted with questions such as, “How can you be consistent in grading performances?” “How can you distinguish between strong and weak oral arguments?” “If your grading is challenged, how will you be able to justify your grades?”

Such concerns may have some merit, but they are generally overstated.<sup>29</sup> Educators in other fields seem to have mastered the skill of constructing rubrics

---

26. See, e.g., James Lindgren & Allison Nagelberg, *Are Scholars Better Teachers*, 73 CHI.-KENT L. REV. 823, 827 (1998) (noting that “[h]ighly cited scholars are perceived by students as better teachers.”).

27. Indeed I view that proposition as a professorial version of the “Lake Wobegon” effect elevated to a “Lake Superior” level! Were it true, one would be compelled to conclude that there are no poor teachers at most highly ranked law schools where faculty are incredibly productive scholars. With all due respect to the many great scholars I know, that proposition does not seem credible. The “Lake Wobegon” effect refers to the tendency of people to believe “that their skills and abilities are above average.” Justin Kruger, *Lake Wobegon Be Gone! The “Below-Average Effect” and the Egocentric Nature of Comparative Ability Judgments*, 77 J. PERSONALITY & SOC. PSYCHOL. 221, 221 (1999).

28. See Deborah Jones Merritt, *Research and Teaching on Law Faculties: An Empirical Exploration*, 73 CHI.-KENT L. REV. 765, 807 (1998) (concluding that no evidence exists of a correlation “between teaching excellence and scholarly productivity”). One problem that the prolific scholar can face is that her students are novices in the field while she is increasingly an expert. Targeting her teaching to help novices becomes more difficult as the gap between her knowledge and that of her students widens. An expert teacher can overcome that problem, but doing so takes time, time that she may be tempted to devote to other endeavors, including additional scholarship.

29. I may be guilty of being much too generous here. Educators in others fields seem to have mastered the skill of constructing rubrics to assess student performances.

to assess student performances.<sup>30</sup> Perhaps we shy away from grading performances because we are unfamiliar with how to construct appropriate rubrics for grading such performances. Possibly we prefer not to acknowledge the fact that skill performances can be graded because we subconsciously realize that once we begin to evaluate student skill performances on anything other than a pass-fail scale, we must also acknowledge that our teaching skills are likewise capable of being graded, or at least they should be. The prospect of being graded for our teaching is not pleasant.<sup>31</sup>

Even less appealing is the recognition that switching to a learning assessment-based model would make intra-faculty comparisons of teaching effectiveness possible. If we were to look at learning outcomes rather than teaching inputs, we would be forced to admit that traditional explanations for differing levels of performance in similar classes (e.g., “I am more demanding.” “She grades more leniently.” “He caters to the students.”) carry no weight.<sup>32</sup> Faculty members who might try to justify the results by explaining that they had loftier goals for students would be compelled to explain what those goals are and why those goals are appropriate for the particular course. To most of us in the professoriate who have experienced success in our writing, whether on law school examinations or in law review articles, the prospect of subjecting ourselves to a system that might actually grade our teaching is not attractive. If we abide by the secret agreement of the professoriate that all tenured faculty members and virtually all untenured faculty members are better than average teachers, we can more easily maintain gentility of the faculty hallway.<sup>33</sup>

I could continue on at length in this essay to further explain why we law professors have a vested interest in maintaining the present state of affairs, but doing so would distract further from the simple point I started out to make. Put simply, the focus on the teacher implied by the phrase “teaching evaluation” has a pernicious effect on legal education. By focusing our attention on teaching evaluations rather than on student learning evaluations, we are unwittingly undercutting our efforts to improve legal education. What we really should care about is whether our students are learning what we want them to learn. By directing our focus on teaching, we tend to evaluate teaching out of context. Good teaching thus becomes a thing unto itself, regardless of how effective that teaching is. Evaluating teaching in the abstract leads us into the thicket of never ending debates over how to define good teaching.

We ought not to be concerned about good teaching, at least not in the abstract. Great teaching performances are of little value if they do not result in

---

30. For a thorough discussion of the use of rubrics in grading, see generally Sophie M. Sparrow, *Describing the Ball: Improve Teaching by Using Rubrics—Explicit Grading Criteria*, 2004 MICH. ST. L. REV. 1.

31. No one, it seems, wants to confront the possibility that so-and-so might actually not be effective as a comfortable belief about our collective competence.

32. I am assuming that no school would accept the faculty member’s own assignment of grades as evidence of effective outcomes.

33. Many law professors also value congeniality highly; they would be reluctant to deprive their colleagues of the label “good teacher.”

effective student learning. If we substitute *learning assessments* for *teaching evaluations*, we would be more likely to develop and use tools for assessing the effectiveness of our teaching.<sup>34</sup> The label itself would drive us to focus on outcomes rather than inputs. Admittedly, one might still elect to focus on evaluating teaching (and then declare doing so to be impossible), but we should not ignore the power of the words we use. If we start focusing on how much our students are learning rather than on how well we are teaching, we likely will find ourselves more attentive to the development of better learning materials going on all around us.<sup>35</sup>

Consider how the simple change from asking about the quality of teaching to asking about extent of student learning might affect the education process. We now assess a faculty member's teaching ability in at least three contexts: when we evaluate a faculty member for retention, promotion, or tenure; when we ask students to evaluate the teacher's performance in a course; and when we determine a faculty member's merit salary increase. So long as teaching quality is the issue, a faculty retention, promotion, and tenure committee can visit a class, skim through the student evaluations, and pass judgment. If, on the other hand, the level of student learning were the issue, the committee would be hard pressed to explain that it had concluded there was significant student learning occurring in a professor's course after its members simply had sat in that professor's class for a short period of time and had looked at evaluations completed by students well before the end of the course. To be true to their charge, the committee members would be compelled to look at the professor's course objectives. They would want to see the assessment tools used by the professor, and some evidence on how students had performed on those assessment tools. If the committee members do anything less, they might find their recommendation rejected by their dean or the provost for lack of substantiation.

Equally important, because the burden of proving merit usually falls on the faculty members under review, they too would feel the need to perfect ways to assess student learning. Faculty members undergoing review would realize that relying on a fistful of non-controversial student evaluation forms and vague descriptions of their skills as lecturers or Socratic dialoguers would be risky business (and we law professors are notoriously risk averse!). They would suddenly have incentives to develop valid learning objectives. They would look for ways to demonstrate student learning. They likely would abandon the dominant practice of relying solely on single, end-of-the-semester examinations

---

34. If we believe CAFE targets will induce the automobile industry to improve automobile mileage efficiency, we should expect teaching effectiveness to do the same for our profession. For a discussion of CAFE targets, see BD. ON ENERGY AND ENVTL. SYS. ET AL., EFFECTIVENESS AND IMPACT OF CORPORATE AVERAGE FUEL ECONOMY (CAFE) STANDARDS 13-19 (2002), available at <http://www.nap.edu/openbook.php?isbn=0309076013>.

35. See, e.g., GERALD F. HESS & STEVEN FRIEDLAND, TECHNIQUES FOR TEACHING LAW 261-64 (1999); GREGORY S. MUNRO, OUTCOMES ASSESSMENT FOR LAW SCHOOLS 11-12 (2000); MICHAEL HUNTER SCHWARTZ, EXPERT LEARNING FOR LAW STUDENTS 16 (2005) (discussing the advantages of the Socratic method); STUCKEY, *supra* note 1, at 123; TEACHING THE LAW SCHOOL CURRICULUM 7 (Steven Friedland & Gerald F. Hess eds., 2004).

to assess student learning. More and more faculty members would start to use interim assessments, including formative evaluations designed to enable students to identify their deficiencies and shore up their weaknesses. Otherwise their students would not have opportunities to improve their performances. Suddenly, faculty members would have real incentives to participate actively in the teaching development sessions offered by the law school and other organizations. And, importantly, having amassed such evidence, faculty members would be likely to insist that the law school acknowledge and reward effective teaching.

Finally, students asked to evaluate their own learning rather than the faculty member's teaching would be less likely to ignore their own role in the learning process. They would be more likely to insist that they be able to understand what it is that they are supposed to be learning and hence to ask for learning objectives. They would also be more likely to demand that they be able to delay undertaking the evaluation until after they had time to assess the congruence between the course learning objectives and what they actually learned. With the focus shifted to them and their learning, students are more likely to become active partners in their own educational experience.

Admittedly, as history has taught us, attempts to improve legal education are never certain to bear fruit. To succeed, we will need to change the law school culture. We will also need to adopt new evaluation techniques. Law schools, however, have not demonstrated a hunger for change. Law professors have not demonstrated a proclivity to be early adopters either. Indeed, we can be assured that any proposed change that might shake the comfortable myths that no one can determine what good teaching is and that most, if not all, law school teachers are equally effective teachers will not receive a warm welcome. Change, however, is necessary. By shifting our attention to assessing student learning rather than continuing to focus on evaluating a professor's teaching, we would create incentives for faculty members to develop their own effectiveness as teachers and thus improve the prospects for change. Therefore, as I indicated at the outset of this essay, I propose heresy. If we are serious about improving legal education we should do the unthinkable. We should abolish the use of teaching evaluations entirely.