

The Compounding Effects of Assessment: How Our Failure to Coordinate Formative Assessments May Impact Their Validity

Lindsey P. Gustafson*

With an ABA-mandated increase in formative assessments,¹ law faculties teaching first-year students are contributing to master calendars to coordinate major assessments.² But contributing to a master calendar alone will not ensure that our entire assessment system, as experienced by first-year students, is valid and improves learning.

A valid assessment is one that accurately measures the learning goals it is designed to measure and justifies the inferences made from it about student learning.³ The data teachers and students receive from formative assessments is helpful only when it is valid, and data from a single formative assessment may lack validity when we fail to account for how students may be juggling multiple assessments from all of their classes.⁴ And—in an effort to improve student learning—we are adding more assessments. The middle of the semesters, which have traditionally been the playground for the Socratic method and for legal writing assignments, may now be filled with a variety of assessment activities,

***Lindsey P. Gustafson**, Professor of Law, University of Arkansas at Little Rock, William H. Bowen School of Law. Special thanks to Michael Hunter Schwartz for his excellent mentorship and to Marci Houseman, one of the most dedicated educators I know. This article was written with the generous support of a writing grant from the Bowen School of Law.

1 See ABA Standards and Rules of Procedure for Approval of Law Schools 2020-2021: ABA STANDARD 314, AMERICAN BAR ASSOCIATION, https://www.americanbar.org/content/dam/aba/administrative/legal_education_and_admissions_to_the_bar/standards/2020-2021/2020-21-aba-standards-and-rules-chapter3.pdf (“A law school shall utilize both formative and summative assessment methods in its curriculum to measure and improve student learning and provide meaningful feedback to students.”).

2 See generally Sandra L. Simpson, *Coordinating Formative Assessment Across the Curriculum: A View from the Associate Dean's Desk*, 95 U. DET. MERCY L. REV. 91 (2017).

3 See THE INTERNATIONAL ENCYCLOPEDIA OF EDUCATIONAL EVALUATION 251 (Herbert J. Walberg & Geneva D. Haertel eds., 1990).

4 See David Boud, *Assessment and Learning: Contradictory or Complementary?*, in ASSESSMENT FOR LEARNING IN HIGHER EDUCATION 38 (Peter Knight ed., 1995) (“Very little attention has been given to the compounding effects of assessment even when we know that it is the total array of demands in a given period which influences how each one is tackled.”).

and some of them may dominate students' time in a way that impacts students' learning in other classes.⁵ Consequently, we must all consider our assessments as students experience them, "holistically and interactively."⁶

Understanding the role our assessment plays in a student's entire menu of assessments requires an understanding, first, of the purposes and challenges of formative assessment and, second, of the demands of a validity analysis, including a consequential validity analysis, which measures the "washback" effect of the assessment on student learning in other courses.⁷ We must ask whether our assessment actually improves student learning, or whether we are simply assuming that more formative assessment must be a good thing.

Without a scholarly understanding of the consequential validity of our formative assessments, we will continue to build assessment systems iteratively, making mistakes and correcting them. But students cannot escape the mistakes we make in our assessments.⁸ The way we assess our students determines how they learn the material, how they will be evaluated by potential employers, and how they will view themselves as students and future attorneys. Our students will benefit from a more vibrant, diverse culture of assessment when we educate ourselves to avoid—rather than just repair—assessment design and coordination mistakes and to broaden our conception of what is possible for student learning.⁹

This article is intended to guide law faculties as we work to create a diverse, coordinated culture of assessment that improves student learning and

5 See Lorrie A. Shepard, *Evaluating Test Validity*, in 19 REVIEW OF RESEARCH IN EDUCATION 405, 408 (Linda Darling-Hammond ed., 1993) ("Often, the use of a test is defended on the grounds that some information (i.e., test scores) is better than none. But for tests with unknown or inadequate validity evidence, this defense can no longer be accepted routinely").

6 See Harry Torrance, *Formative assessment at the crossroads: conformative, deformative and transformative assessment*, 38 OXFORD REV. EDUC. 323, 334 (2012) (noting that "assessment is always formative, but not necessarily in a positive way").

7 See Samuel Messick, *Validity and Washback in Language Testing*, EMOT. TESTING SERVS. (1996) (connecting "washback," a term used in language instruction, to validity inquiries).

8 See Katrien Struyven et al., *Students' perceptions about assessment in higher education: a review* 10 (paper presented at the Joint Northumbria/Earli SIG Assessment and Evaluation Conference: Learning Communities and Assessment Cultures, University of Northumbria at Newcastle, Longhirst Campus, Aug. 28-30, 2002), <http://www.leeds.ac.uk/educol/documents/00002255.htm> ("[T]he most influential feature of the learning environment is the nature of the assessment procedures."); Boud, *supra* note 4, at 35 (noting the effects of bad assessment "are far more potent than they are for any other aspect of teaching").

9 See Cassandra L. Hill, *The Elephant in the Law School Assessment Room: The Role of Student Responsibility and Motivating Our Students to Learn*, 56 HOW. L.J. 447, 453 (2013) ("Since law professors are not required to have formal teaching credentials or an advanced degree in education to teach in law school, they must be proactive about improving their teaching skills and placing student learning at the forefront of their teaching."); Susan Hanley Duncan, *The New Accreditation Standards Are Coming to a Law School Near You—What You Need to Know About Learning Outcomes and Assessment*, 16 J. LEG. WRITING 605 (2010) (recommending that law faculty begin to consider whether institutional leaders support and value a culture of assessment).

generates data with validity that will inform our students, our teaching, and also future employers. After this article defines and describes the promise and complexity of formative assessment in Part I, it provides evidence in Part II that a failure to create a deeply coordinated formative assessment system will lessen the validity of the entire system. Finally, Part III provides best practices for creating a valid, deeply coordinated formative assessment system. When we each create assessments that justify the use of students' time, that allow for transparent learning linked with clear learning goals, and that encourage student autonomy and confidence, we are more likely to design assessments that will improve rather than detract from students' learning in other courses.

I. The Promise and Complexity of Formative Assessment

The limited data we have from law schools on the impact of more formative assessment supports the claim that increased formative assessment in a single course will improve students' final exam performances.¹⁰ Critically, two recent studies have shown that students who received individualized feedback during the semester outperformed the students who did not on final exams, not just in the class where they received the feedback, but in every single class they were taking.¹¹ One study's authors note the "likelihood of this occurring by chance is one in 256."¹²

At the base of any expansion of formative assessment, rooting it and driving it, should be our understanding of why we assess, how we should assess, and how best to determine whether our assessment furthers student learning.¹³ Without this understanding, we risk finding out the hard way "how construct labels can smuggle in whole theories without [our] being aware of the choices [we] have made."¹⁴ Assessment is not formative only when it has a positive

10 See Daniel Schwarcz & Dion Farganis, *The Impact of Individualized Feedback on Law Student Performance*, 67 J. LEGAL EDUC. 139, 142 (2017) (finding that formative assessment improved performance on final exams for students with below-median entering credentials); Deborah Jones Merritt et al., *Formative Assessments: A Law School Case Study*, 94 U. DET. MERCY L. REV. 387 (2017) (finding the same); Carol Springer Sargent & Andrea A. Curcio, *Empirical Evidence That Formative Assessments Improve Final Exams*, 61 J. LEGAL EDUC. 379, 383-84 (2012) (finding that formative assessment improved performance on final exams for students with above-median entering credentials); Andrea A. Curcio et al., *Developing an Empirical Model to Test Whether Required Writing Exercises or Other Changes in Large-Section Law Class Teaching Methodologies Result in Improved Exam Performance*, 57 J. LEGAL EDUC. 195, 197 (2007) (finding the same); Andrea A. Curcio et al., *Does Practice Make Perfect? An Empirical Examination of the Impact of Practice Essays on Essay Exam Performance*, 35 FLA. ST. U. L. REV. 271, 280-82, 302-06 (2008) (finding the same).

11 Schwarcz & Farganis, *supra* note 10, at 142.

12 *Id.*

13 See Boud, *supra* note 4, at 36-37 ("Assessment always leads to learning. But the fundamental question is, 'what kind of learning?' What do our acts of assessment communicate to students?").

14 Shepard, *supra* note 5, at 425; see also Mantz Yorke, *Formative Assessment in Higher Education: Moves Towards Theory and the Enhancement of Pedagogic Practice*, 45 HIGHER EDUC. 477, 484 (2003) ("Untheorised assessment (as is widely used in higher education) increases the risk of

impact on learning; students are impacted by everything they hear and experience in our course. One teaching and learning expert warned that the assessment practices in higher education generally “act[] as a mechanism to control students that is far more pervasive and insidious than most staff would be prepared to acknowledge. It appears to conceal the deficiencies of teaching as much as it does to promote learning.”¹⁵ To expand our use of formative assessment, we must understand its definition, its complexities, how to make it valid, and how to make it consequentially valid for students’ learning generally.

A. *The Complexities of Formative Assessment*

Whereas summative assessment is given at the end of a unit of study to evaluate and rank students’ performances, formative assessment is given during learning to improve that learning.¹⁶ It is “quintessentially process-oriented”¹⁷: It is formative only if it is used to alter the process of learning and teaching.¹⁸ Because formative assessment is dialogic, it is most successful when the teacher and the student fully understand each other and collaborate to create a best performance.¹⁹ An accurate productive dialogue between teacher and student is arguably most achievable in higher education where the students are adults, they are voluntarily engaging in the learning, and they are likely experienced learners.²⁰

Because formative assessment works only when the teacher and the student come together in a shared communication to move the student from the known to the unknown, designing and executing formative assessment demands more

partiality [T]heorisation is needed if some important aspects of assessment are not to be marginalised.”).

- 15 Boud, *supra* note 4, at 35 (“There is probably more bad practice and ignorance of significant issues in the area of assessment than in any other aspect of higher education.”); see Margaret Bearman et al., *How university teachers design assessments: a cross-disciplinary study*, 74 HIGHER EDUC. 49, 50 (2016), <https://link.springer.com/article/10.1007/s10734-016-0027-7> (noting studies that report the gap between “educators’ aims for their assessments and how they are actually implemented in practice”).
- 16 Olympia Duhart, *The “F” Word: The Top Five Complaints (and Solutions) About Formative Assessment*, 67 J. LEGAL EDUC. 531 (2018)
- 17 Yorke, *supra* note 14, at 486.
- 18 See Elizabeth M. Bloom, *A Law School Game Changer: (Trans)formative Feedback*, 41 OHIO N.U. L. REV. 227, 232 (2015) (“Assessment is not formative unless the instructor obtains the information and uses it to adapt instruction to meet the needs of students. In fact, [w]hat makes formative assessment *formative* is that it is immediately used to make adjustments so as to *form* new learning.”).
- 19 See Yorke, *supra* note 14, at 485 (“The exchanges between teacher and student are . . . mutually hermeneutic, in that each is seeking to interpret and understand the communications of the other with the aim that the student will become better equipped to deal with future challenges of varying kinds.”).
- 20 See Torrance, *supra* note 6, at 328.

skill than designing and executing a summative assessment, which focuses only on evaluating and ranking students.²¹

The complexity of formative assessment may be obscured by the “closing the gap” metaphors used to describe its goal: Formative assessment is said to be “emancipatory” by moving students “beyond the current boundaries of knowledge.”²² The teacher uses the assessment to identify a student’s “zone of proximal development,” which describes the area between a student’s existing problem-solving ability and his or her potential ability given guidance and support from a teacher.²³ Most learning, however, requires more than simply transferring knowledge from a teacher to a student; both teachers and students must probe and understand all of the gaps that exist until they have a shared understanding.²⁴ Learning often happens in a loop, with assessments articulating our learning goals and becoming part of a profile of demands to which students should respond, and then with student responses coming back to us for our feedback.²⁵

The profile of demands to which students must respond includes formative *assessment for learning*, a term used to describe assessments that provide feedback on student learning in the moment, assessment that may be happening daily in the classroom.²⁶ When it is done well, this kind of feedback informs both the teacher and the student while learning is happening, and works to help students become more aware of their learning and better self-assessors.²⁷ The Socratic method may work as this kind of assessment, although while other students may learn by following the questions and answers, the teacher can

21 *See id.* at 484.

22 Yorke, *supra* note 14, at 477-78.

23 *Id.* at 478.

24 *See* Torrance, *supra* note 6, at 333-34 (criticizing the “building block” view of knowledge).

25 *See* NELSON MILLER, *TEACHING LAW: A FRAMEWORK FOR INSTRUCTIONAL MASTERY* 100-01 (2010); David Boud & Elizabeth Molloy, *Rethinking models of feedback for learning: the challenge of design*, 38 *ASSESSMENT & EVAL. IN HIGHER EDUC.* 698, 701 (2013) (noting that without a feedback loop, “teachers are blind to the consequences of their actions and cannot therefore act effectively to improve the quality of learning”); Rogelio A. Lasso, *Is Our Students Learning? Using Assessments to Measure and Improve Law School Learning and Performance*, 15 *BARRY L. REV.* 73, 73 (2010).

26 *See* L. DEE FINK, *CREATING SIGNIFICANT LEARNING EXPERIENCES: AN INTEGRATED APPROACH TO DESIGNING COLLEGE COURSES* 21 (2003) (describing classroom assessment techniques where feedback, usually ungraded, is frequently provided to students to enhance the quality of their learning); A. Wade Boykin, *Human Diversity, Assessment in Education and the Achievement of Excellence and Equity*, 83 *J. NEGRO EDUC.* 499, 506 (2014) (citing the recommendation of the National Research Council in 2001 to focus less on assessment of learning and more on assessment for learning); *see also* Torrance, *supra* note 6, at 335 (opining that “the various theoretical tributaries which have fed contemporary formative assessment debates and practices have largely merged into a rather undifferentiated commitment to the development of ‘assessment for learning’”).

27 *See* Boykin, *supra* note 26, at 506-07 (describing the extensive scholarship of Rick Stiggins on assessment for learning).

assess only the student responding to the specific question.²⁸ And even the assessment of that single student is limited because it involves assessment conditions—being called on in front of peers for an oral response—that are typically not part of a final assessment.

Because of the complexity of formative assessment, even in higher education with experienced adult students, formative assessment completed without an understanding of its purpose and its best practices may not fully realize its potential.²⁹ No matter how well intentioned, it may even fail to be a valid assessment.

B. Measuring the Validity of Assessment

When an assessment fails to detect students' true learning, that assessment is not valid.³⁰ An assessment is valid when it is coordinated with learning goals it is intended to assess,³¹ when the assessment actually measures the learning goals as intended, and when the inferences about student learning drawn from the assessment are justified.³²

The latter half of this definition is deceptively simple but key: The inferences drawn from the assessment, and not just the assessment itself, must be valid.³³ It is not enough to establish the validity of an assessment within the confines of an assessment itself.³⁴ The claims we make about and the actions we take because of our assessments must be based on evidence supporting those inferences, not on theory or logic alone.³⁵ Although formative assessment paired with feedback logically should improve student achievement, to prove its validity we must go beyond “a rather nebulous” assertion that formative

28 See Elizabeth G. Porter, *The Socratic Method*, in *BUILDING ON BEST PRACTICES: TRANSFORMING LEGAL EDUCATION IN A CHANGING WORLD* 101 (Deborah Maranville et al. eds., 2015) (detailing the “rates of perceived mental exertion” in different iterations of the Socratic Method).

29 *Id.*

30 See Ron M. Aizen, *Four Ways to Better 1L Assessments*, 54 *DUKE L.J.* 765, 773 (2004) (defining a valid assessment as one that “accurately reflect[s] the knowledge, ability or other construct” that it is designed to measure). In contrast, an assessment is reliable “to the extent [it has an] acceptable consistency or accuracy for a particular use.” WYNNE HARLEN, *ASSESSMENT OF LEARNING* 22 (2007).

31 HARLEN, *supra* note 30, at 21.

32 INTERNATIONAL ENCYCLOPEDIA OF EDUCATIONAL EVALUATION, *supra* note 3, at 251.

33 “[T]he validation of inferences using tests rather than just tests is a very important idea.” Donald B. Rubin, *On Messick and the Validity of Inferences*, in *TEST VALIDITY* 242 (Howard Wainer & Henry I. Braun eds., 1988).

34 See Shepard, *supra* note 5, at 447 (emphasizing that “[o]ne validates, not a test, but an interpretation of data arising from a specified procedure”); William H. Angoff, *Validity: An Evolving Concept*, in *TEST VALIDITY*, *supra* note 34, at 29 (“In the course of these passing decades it has become clear that it was the subject’s responses to the test, even more, the inferences and interpretations to be drawn from the responses, that were to be validated.”).

35 See Shepard, *supra* note 5, at 406–08.

assessment is good and show a link between that assessment and student achievement.³⁶

Validation efforts therefore begin with a hypothesis as to what the assessment will reveal or how it will impact students' learning, but that hypothesis should be challenged by ensuring there is no evidence to the contrary.³⁷ Proving validity is a complex, continuing process that is daunting even for experienced psychometricians; even they have been accused of "rak[ing] together miscellaneous correlations" and thereby offering "validity evidence in practice that is simplistic and incomplete."³⁸ Experts and the rest of us alike may be giving only a "ritualistic recitation" of the requirements of validity, "rather than a serious investigation."³⁹ Validity evidence may often be "simplistic and incomplete," with "miscellaneous correlations" used as evidence of validity.⁴⁰

A classic example among psychometricians of the demands of validity is the use of a school readiness test to measure which children are ready for kindergarten and which would benefit from waiting a year.⁴¹ Proving the validity of the readiness test would require showing a correlation between test scores and school performance, that students who waited a year performed better (for any number of reasons) than they would have without the extra year.⁴² But these tests continue to be administered without this proof, even in the face of some studies that show no academic benefit and some emotional harm from the extra year.⁴³ Providing proof of validity is daunting, but proving the validity of an assessment is critical to justifying that assessment.⁴⁴

A law school setting provides a particularly difficult space for gathering validity evidence. Most of us teach as we were taught, reacting to successes and failures as we strive to develop into better teachers class by class and student by student, and we are not trained psychometricians. Validity evidence we can obtain in a law school setting is often only through "natural experiments."⁴⁵

C. Measuring the Consequential Validity of Assessments: Washback on Learning

In 1989, Samuel Messick directed validity inquiries beyond a consideration of the actual assessment and its use to both the intended and unintended social

³⁶ Torrance, *supra* note 6, at 327.

³⁷ Shepard, *supra* note 5, at 407.

³⁸ *Id.*

³⁹ Angoff, *supra* note 34, at 30.

⁴⁰ Shepard, *supra* note 5, at 407.

⁴¹ *See id.*

⁴² *Id.*

⁴³ *Id.*

⁴⁴ Angoff, *supra* note 34, at 19.

⁴⁵ *See* studies described *supra* note 10.

and learning consequences of an assessment—to ensure the *consequential validity* of an assessment.⁴⁶ According to Messick’s definition, consequential validity is high when an assessment has a “positive washback effect on learning” and low when it encourages (either intentionally or unintentionally) methods of learning counter to what is desired.⁴⁷ Messick warned, for example, that it “should not be taken for granted that richly contextualized assessment tasks are uniformly good for all students [because] contextual features that engage and motivate one student and facilitate effective task functioning may alienate and confuse another student and bias or distort task functioning.”⁴⁸

Some psychometricians argue that considering consequences as part of a validity determination results in confusion rather than clarity and ties a validity inquiry to factors outside the assessment creator’s control.⁴⁹ In her influential writings on validation, Professor Lorrie Shepard acknowledged it is not always possible to predict consequences or to identify a cause-and-effect link.⁵⁰ Students may react with a negative attitude toward an assessment for any number of reasons, including reasons that are difficult to predict or control.⁵¹ Further, the full consequence of an assessment may not be apparent during the planning stages or even after the assessment has been administered. She instead recommends focusing on a core question: Considering all the assessment is designed to do, and all the evidence of the impact it may have on student learning, are students better with the assessment or without it?⁵²

D. The Challenges of Consequential Validity Analyses in Law School

An analysis of whether students are better off with an assessment requires that we consider the entire landscape of assessment students’ face. Even when we synchronize our own assessments and ground them in learning theory, students’ learning may be cluttered or confused by other responsibilities or

46 See Samuel Messick, *Validity*, in EDUCATIONAL MEASUREMENT 13, 17 (Robert L. Linn ed., 3d ed., 1989).

47 *Id.*; see Boud, *supra* note 4, at 35; Boykin, *supra* note 26, at 509 (“Consequential validity concerns the individual, social, or societal ramifications of the deployment of an assessment, which are deemed either harmful or beneficial.”).

48 Samuel Messick, *The Interplay of Evidence and Consequences in the Validation of Performance Assessments*, EDUC. TESTING SERV. 25 (1992).

49 See W. JAMES POPHAM, CLASSROOM ASSESSMENT: WHAT TEACHERS NEED TO KNOW 118 (8th ed. 2017) (“the notion of consequential validity is apt to confuse the central thrust of validity—namely to confirm or disconfirm the defensibility of the score-based inferences we make about our students”). In large part, this concern came from developers of standardized tests, who claimed they could not and should not have their tests’ validity determined by their uses and interpretations, which they could not control. See Mark D. Reckase, *Consequential Validity from the Test Developer’s Perspective*, 17 EDUC. MEASUREMENTS: ISSUES AND PRAC. 13 (2005).

50 See Shepard, *supra* note 5, at 411.

51 See *id.*

52 *Id.* at 408.

experiences.⁵³ Students are more likely to give less attention to an assessment if it is “located among a thicket of examinations.”⁵⁴ Because an assessment to be valid must give us accurate data on student learning, an assessment lacks validity when its location among other assessments may “cloud, misrepresent, or even fail to detect the ability, competence, or potential of many students.”⁵⁵

A consequential validity analysis of assessments in the first year of law school may be easier than similar analyses done in higher education because generally all of our students are in the same classes and thus have the same set of assessments. To understand the combined assessment pressures on our students, we need coordinate with only a limited number of colleagues. The conditions at law schools as institutions, however, provide for the perfect storm of uncoordinated assessment activities that may together lack consequential validity. First, as an academy we lack a tradition of formative assessment beyond the Socratic method, so we are forced to borrow techniques from other disciplines. When we borrow techniques, we may take them without understanding the scholarship behind them, tolerating a deficit of understanding in assessment techniques that we would never accept in our scholarship.⁵⁶ Second, any level of coordination between courses is relatively new to the academy. Law professors traditionally act independently within a siloed curriculum. We teach how we teach, often without consulting other faculty members. Finally, we expect law school to be competitive and challenging for students, and we demand students build resilience, time-management skills, and grit. We may therefore be less likely to attribute student workload struggles with a flaw in our assessment system. And the students most likely to be impacted by poorly coordinated or considered workloads are the struggling students, especially those with maladaptive responses to feedback—the very students we would like to inform and empower with our formative assessments.

53 See Struyven et al., *supra* note 8, at 24; Boud, *supra* note 4, at 38 (noting that students’ learning is most influenced by “how the student *interprets* the task at hand and the context of the assessment”). Taken to an extreme, this makes assessment unreasonably situational because “the outcomes of assessment as ‘lived’ by students are never entirely predictable, and the quest for a ‘perfect’ system of assessment is, in one sense, doomed from the outset.” Struyven et al., *supra* note 8, at 24. But the learner is key. “The learner is an active partner in the process of learning, teaching, and assessment. S/he selects, perceives, interprets, and integrates new information to form a coherent and meaningful whole with her/his prior knowledge and former experiences.” *Id.*

54 Boud, *supra* note 4, at 38.

55 Boykin, *supra* note 26, at 510 (writing specifically on the impact of uncoordinated assessment practices on students from diverse backgrounds).

56 See David Boud, *Assessment and the Promotion of Academic Values*, 15 *STUDIES IN HIGHER EDUC.* 101 (1990) (noting that “many current assessment practices are incompatible with the goals of independence, thoughtfulness and critical analysis to which most academics would subscribe”); Linda S. Anderson, *Incorporating Adult Learning Theory into Law School Classrooms: Small Steps Leading to Large Results*, *APPALACHIAN J. L.* 127 (2006) (noting that educators should learn the new technique and the why behind that technique).

II. Evidence that Failure to Create a Deeply Coordinated Formative Assessment System Lessens Its Validity

Even after Samuel Messick introduced the concept of consequential validity and directed attention toward an assessment's impact on learning, "[v]ery little attention has been given to the compounding effects of assessment."⁵⁷ Even less attention has been given to analyzing whether attempts at coordinating multiple formative assessments in higher education improves learning.⁵⁸

But two studies, described below, are relevant both to the compounding effects of assessment in higher education and to the effective coordination of those assessments: First, a 2010 study of law school students' assessment preferences found that while at the beginning of law school a majority of students prefer multiple assessments, by the end of the first semester a majority of students do not. Second, a 2008 study of the assessment system instituted at the competitive Lund University engineering program concluded that, although the assessment system was coordinated across the curriculum from the instructors' perspective, it was not correlated to student learning. Combined, these studies indicate that creating a coordinated assessment system that engages students and encourages student learning requires we do more than simply ensure none of our assessments conflict on a calendar.

A. A Study of Student Preferences for Formative Assessment

In 2010, Emily Zimmerman reported the results from her survey of two law school classes on their assessment practices.⁵⁹ She surveyed students at the beginning and again at the end of the year. Her results show that while at the beginning of law school a majority of students (roughly seventy-three percent) agreed that they prefer multiple graded assessments throughout a course, by the end of the first year the number of students still agreeing dropped to roughly fifty-seven percent.⁶⁰ Student preferences also declined for multiple ungraded assignments, from roughly forty-seven percent on the entrance survey to roughly twenty-five percent on the exit survey.⁶¹

Professor Zimmerman acknowledges the limits of her data, and calls for more research into students' assessment preferences and the reasons for those preferences.⁶² Students are important stakeholders in assessment, but they

57 Boud, *supra* note 4, at 35.

58 See Emily Zimmerman, *What Do Law Students Want?: The Missing Piece of the Assessment Puzzle*, 42 RUTGERS L.J. 1 (2010) (indicating little research has been done on how the *frequency* of assessment impacts final exam performance).

59 *Id.*

60 *Id.* at 60.

61 *Id.* at 61 (giving the data for 2007–2008; in 2008–2009, only roughly thirty-nine percent of respondents indicated a preference for multiple ungraded assignments on the entrance survey).

62 *Id.* at 65.

may not be informed on how best to improve their own learning.⁶³ It may be that during the first-year students become acculturated to traditional law school assessment practices—with Socratic method and one summative exam—and begin to trust that structure. It may also be that by the end of their first year students are so traumatized by exams they cannot imagine asking for more of them.

Professor Zimmerman hypothesizes that students may be responding to workload concerns, which would explain their aversion to ungraded assessments—assessments that do not sound as though they will recognize the students’ efforts with points to raise their grade.⁶⁴ As Professor Zimmerman notes, “[t]he intense nature of the workload in law school and students’ reactions to it may be significant factors in determining the impact of the law school experience upon law students.”⁶⁵ She recommends that before we add assessments, we consider “students’ workload—in particular, first-year students’ workload . . . holistically, and not only on a course-by-course basis.”⁶⁶ Teachers of first-year courses should meet together to coordinate not only scheduling but the methods of assessment to decide whether “students are being expected to complete such a large number of assignments over the course of a term so as to undercut the pedagogical benefits of those assignments.”⁶⁷

Similar concerns were raised in a 2002 study out of the Centre for Research on Teacher and Higher Education at the University of Leuven in Belgium, which described findings from a series of structured group sessions that elicited students’ views on what helped or hindered their development.⁶⁸ Students viewed assessment as a “powerful motivator and a major vehicle for learning.”⁶⁹ But students also acknowledged that under a heavy workload of assessment, they reduced the depth of their understanding and worried that the work would “just wash over students.”⁷⁰ Students reported that although they valued the introduction of alternative assessment methods, often the cluster of assignments meant “they simply did not have the time to invest in this level of learning.”⁷¹

63 *Id.* at 67 n.205.

64 *Id.* at 67.

65 *Id.*

66 *Id.*

67 *Id.* at 68; see also Dalia Bedewy & Adel Gabriel, *Examining perceptions of academic stress and its sources among university students: The Perception of Academic Stress Scale*, HEALTH PSYCHOLOGY OPEN, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5193280> (2015) (citing multiple studies of student stress occurring around “examination periods”). In Part IV (A) below, where I recommend teachers include each assessment’s learning goals when they meet to discuss and coordinate.

68 Struyven et al., *supra* note 8, at 20.

69 *Id.*

70 *Id.*

71 *Id.* at 21.

B. A Study of a Coordinated Formative Assessment System that Became Disconnected from Student Learning

As law faculties are adding more formative assessments, many are coordinating major assessments on a shared calendar, but no law school has undertaken a comprehensive study of its overall assessment systems to determine whether this coordination actually improves the validity of the assessments. Such a study was conducted on the assessment system of the Lund University engineering program in Sweden.⁷² The study is instructive because the learning environment there is similar to the learning environment in most law schools, with students in the program facing high educational demands from traditionally siloed instructors. Further, instructors in the engineering program were motivated to adopt a range of assessments because they wanted to improve student learning broadly and to reduce the number of failing students. After instituting the assessment system and coordinating it across the curriculum, the program invited assessment experts from outside the school to determine whether the system was actually improving student learning.⁷³

The study's purpose was to analyze the entire assessment process's impact as a social practice in context, so as not simply to assume certain outcomes for student learning based on the coordinated structure of the multiple assessments.⁷⁴ To analyze the broader effects of assessment, the authors monitored fail rates, reviewed all of the syllabi and assessment assignments, and collected data from student questionnaires, institutional reports, focus groups and interviews with students and teachers.⁷⁵

The study's authors described the assessment system as "tight and complicated," but not clearly coordinated among courses.⁷⁶ The expanded assessment system was intended to "help more students pass . . . by distributing their study efforts more evenly over time and between their different courses."⁷⁷ A chart of students' assignments appears coordinated for the students, but the assessments were "unevenly expanded": Some courses had mandatory weekly assessments, others had a midterm and a final, and others had only a final exam.⁷⁸

72 Asa Lindberg-Sand & Thomas Olsson, *Sustainable Assessment? Critical Features of the Assessment Process in a Modularised Engineering Programme*, 47 INT'L J. OF EDUC. RESEARCH 165, 169 (2008) (citing a 2005 Swedish report). Professor Lindberg-Sand is with the Lund University Centre for Educational Development and Professor Olsson is a member of the Lund University Engineering faculty.

73 *Id.* at 168.

74 *Id.*

75 *Id.* at 168-69.

76 *Id.* at 170.

77 *Id.* at 172.

78 *Id.* at 170-71.

The study's authors concluded that, rather than helping more students pass the rigorous curriculum, the intricate assessment system had in critical ways become separated from actual student learning.⁷⁹ If students failed to keep up with the planned pace of the "delicate balancing of all the different tasks in relation to each other," they were isolated while the teaching continued on.⁸⁰ The program did not have a mechanism to identify and address the needs of failing students within the flow of the coursework.⁸¹ Struggling students instead just tried working harder and "narrowing their focus to the most immediate events that [lay] ahead."⁸² If this did not work, students would "strategically drop" first non-mandatory assessments and then mandatory assessments that seem less critical.⁸³ In this way, many students built "assessment debts" to be paid back later.⁸⁴ Because of the demanding and competitive culture of the engineering program, students routinely attributed these missed assignments to faults in their own personal habits, not to a poorly designed and executed assessment system.⁸⁵

Faculty likewise failed to view student struggles as evidence of a flawed assessment system and instead blamed the institution for admitting weaker students.⁸⁶ The teachers were able to justify the time demands of their assessment assignments precisely and clearly to the study's authors.⁸⁷ But when the teachers added assignments to the assessment system, "they appear[ed] to get no feedback on how these changes affect[ed] the learning process of the students."⁸⁸ Students were expected to struggle with the challenging curriculum.⁸⁹ And because the curriculum was modularized, faculty often failed to see the consequences of their individual assignments on student's success across the curriculum.⁹⁰ Students continued to attend classes and lab sessions, and appeared to be participating, and therefore learning.⁹¹

Because students did not attribute their failures to a poorly structured assessment system, and because faculty did not realize the assessment system

79 *Id.* at 173.

80 *Id.* at 172.

81 *Id.* at 171.

82 *Id.*

83 *Id.*

84 *Id.* at 172.

85 *Id.* at 171-72.

86 *Id.* at 171.

87 *Id.* at 173.

88 *Id.* at 172.

89 *Id.* at 171.

90 *Id.* at 172.

91 *Id.* at 172-73.

was not actually improving student learning, the flaws in the system were “invisible” to the stakeholders and created “unintended consequences for student learning.”⁹²

Nothing in any of these studies indicates we should not increase the formative assessments of our students; indeed, there is growing evidence that doing so can improve student learning.⁹³ But we cannot simply assume that students will appreciate the benefits of the additional assessments, and we cannot assume that simply coordinating calendars—although this is a logical beginning—is enough to deeply coordinate our assessment systems.

III. Best Practices for Creating Valid, Deeply Coordinated Formative Assessment Systems

Given the evidence that even coordinated assessment systems can become separated from student learning,⁹⁴ how do we deeply coordinate assessments and avoid an assessment system that instead overwhelms students, clutters or fragments their learning, or discourages them early in their first semester? We must coordinate beyond shared calendars, starting in our own classrooms by ensuring that our own assessment activities, as a slice of the student-time pie, are designed with and justified by best practices that encourage that assessment’s consequential validity, including 1) making the assessment system’s alignment with learning goals transparent to students; 2) using rubrics to create a shared language of instruction; 3) ensuring the assessment system encourages student autonomy; 4) setting high expectations and displaying confidence that students can meet those expectations; and 5) regularly reviewing the entire assessment system, paying particular attention to students’ ownership of their own learning within the system. When we ground our formative assessment decisions in best practices, we are better able to communicate our decisions to students, and better able to more deeply coordinate with other faculty members.

A. Make the Assessment System’s Alignment with Learning Goals Transparent

We justify our formative assessment decisions by carefully articulating and connecting the assessment to our course goals.⁹⁵ Without this connection to learning goals, we cannot decide what to assess; how to assess, either formatively, summatively, or both; how much time in and out of class the

92 *Id.* at 172.

93 *See supra* note 10.

94 *See* Lindberg-Sand & Olsson, *supra* note 72, at 173.

95 *See* Barbara Glesner Fines, *Outcomes Assessment for Improving Student Learning*, in BUILDING ON BEST PRACTICES, *supra* note 28, at 94 (labeling the process of identifying course goals as course-level “outcomes assessment”); Okianer Christian Dark, *Statement of Good Practices in Legal Education: Principle 6: Good Practice Communicates High Expectations*, 49 J. LEGAL EDUC. 401 (1999).

assessment should require; and what assessments we can forgo if necessary to ease a student's workload.

The more incremental and connected the learning goals are—short-term goals leading to long-term goals—the more detailed the planning and the coordination may be.⁹⁶ For example, it may be that some of the assessments traditionally performed in a legal writing or other skills course overlap with assessments now assigned in doctrinal courses. A few years ago, when I discovered the civil procedure professor was having her students submit a written brief of a case that she would then review with written feedback, I cut that assignment from my first-semester writing course. We had overlapping learning goals, and the students did not need two written assessments of that goal.

Assessing incremental goals not only improves coordination for faculty members, but also allows students to see their success or failure in specific, more accurate ways, which can improve their efficacy and motivation to keep working.⁹⁷ Efficacy expectations are built by a student's past successes and failures; students who believe that an isolated learning strategy or knowledge, skill, or value has been mastered feel more control over their own learning.⁹⁸ Students with high efficacy expectations are more likely to be engaged, work harder on assigned tasks, and accept more challenges.⁹⁹ Students may be less likely to feel effective and more frustrated with the assessment system generally if the learning goals are not specific enough for a student to accurately interpret and build on the feedback.¹⁰⁰

Studies have also demonstrated that students' motivation increases when students see achievable but challenging learning goals connected to the assessments.¹⁰¹ Learning goals, which are goals to acquire or improve a skill, are more likely to motivate than performance goals, which are goals to document or demonstrate the skill.¹⁰² Most assessments involve performance goals, but many could also include a connection to a broader learning goal, possibly through a rubric used on repeated assessments that tracks improvement

96 E. Scott Fruehwald, *How to Help Students from Disadvantaged Backgrounds Succeed in Law School*, 1 TEX. A&M L. REV. 83, 94-95 (2013).

97 *See id.* at 94 ("Connecting short-term goals with long-term goals can also create positive outcome expectancies").

98 *See* MICHAEL HUNTER SCHWARTZ ET AL., TEACHING LAW BY DESIGN: ENGAGING STUDENTS FROM THE SYLLABUS TO THE FINAL EXAM 63 (2009); *see also infra* section III.C.

99 Fruehwald, *supra* note 96, at 95.

100 *Id.*; *see also* HARLEN, *supra* note 30, at 145 ("The more open we are about assessment procedures the easier it is for students to assess their own work. This is vital for using assessment to help learning.").

101 *See* Fruehwald, *supra* note 96, at 92.

102 *See* Sarah J. Adams-Schoen, *Of Old Dogs and New Tricks—Can Law Schools Really Fix Students' Fixed Mindsets?*, 19 J. LEG. WRITING INST. 3, 19-20 (2014); Patricia A. Smiley & Carol S. Sweck, *Individual Differences in Achievement Goals among Young Children*, 65 CHILD DEV. 1723, 1724-25 (1994).

along a continuum. Students who can see and adopt learning goals are more likely to reach for deeper understanding, to ask for help, to persist, and to be comfortable with challenging tasks.¹⁰³ Students with learning goals are more likely to make better decisions with their time and attention as they develop more broadly as legal problem-solvers.¹⁰⁴

Obviously, students are less likely to benefit from incremental assessment if the students misunderstand the way the smaller learning goals build to the larger learning goals.¹⁰⁵ As an example of the misunderstanding that may occur, in my first-year property course I routinely assess with two kinds of multiple-choice questions: multiple-choice questions that come at the beginning of a unit to measure a student's readiness to learn,¹⁰⁶ and multiple-choice questions that are part of a midterm or final to measure a student's learning in part or all of a course. Because these categories of questions have different learning goals, I did not intend the readiness-assessment multiple-choice questions to prepare students for the kinds of questions they would encounter on the final exam. But I discovered in conferences after the final exam that some students who did well on readiness-assessment quizzes believed they were adequately prepared for the much harder final exam questions. To clarify the differences between the two assessments, I had to explain my assessment system to the students, and I added a review of harder multiple-choice questions to the beginning of each class so students could see the differences between the questions.

But students' time is finite, and we cannot be so incremental in our assessment that students do not have adequate time to complete the assignments. Each student's total diet of assessment matters to that student's learning, and students' perception of their workload impacts their learning. Studies show that when students view an individual assignment or a combined course workload as excessive and unproductive, they are more likely to adopt surface approaches to learning.¹⁰⁷ And if an assessment is not carefully designed to force students to deeper understanding, a teacher may not discover that a student's high score does not reflect true understanding of the material.¹⁰⁸ In this way, the assessment has a low validity in a traditional sense because

103 Fruehwald, *supra* note 96, at 92.

104 *Id.*

105 See HESS ET AL., *TECHNIQUES FOR TEACHING LAW* 2,261-62 (2011) (students should know how the assessment forwards their learning and prepares them for graded assignments).

106 These multiple-choice questions are given as part of team-based learning. See Sophie M. Sparrow & Margaret Sova McCabe, *Team-Based Learning in Law*, 18 *LEGAL WRITING: J. LEGAL WRITING INST.* 153 (2012).

107 See Struyven, *supra* note 8, at 20.

108 *Id.*

the inferences the teacher draws from student performance are not accurate, appropriate, or useful.¹⁰⁹

We must also avoid structuring students' learning so much that we become mechanistic, rather than interactive, in our assessment.¹¹⁰ Explicit, shared learning goals inform the learning, but the experience with students in and out of the classroom should improve the quality and diversity of our learning goals.¹¹¹ In other words, "Are we trying to get students to jump through pre-specified hoops, but making the nature of those hoops more apparent and encouraging students to better understand how the objectives of a course can be met; or are we trying to get students to think for themselves?"¹¹² Learning goals should be flexible, and assessment should be interactive, if we are to balance our need for structured learning goals against the need to teach the students in front of us and to promote their own autonomy in learning.¹¹³

We must clearly communicate and then listen to our students. Assessments will have a low validity if students do not see what we see. When students misunderstand the purpose of the assessments and view the time spent on them as excessive or wasted, they are less likely to learn from that assessment (or just from us, if their frustration reaches that level).¹¹⁴ The more transparent we are with how our assessment decisions align with learning goals, the more likely students will be to exert the expected effort and time on the assessment, and the more likely the results will be accurately interpreted.¹¹⁵ In short, "we teach better when we have identified a course's goals, and students learn better when they understand what these goals are."¹¹⁶

109 See Jose-Luis Menendez-Varela & Evan Gregori-Giralt, *The contribution of rubrics to the validity of performance assessment: a study of the conservation-restoration and design undergraduate degrees*, 41 *ASSESSMENT & EVAL. HIGHER EDUC.* 228, 230 (2015) ("Validity is an 'adequately grounded judgment that learning achievements correspond to learning goals[.]'").

110 Torrance, *supra* note 6, at 328-29.

111 *Id.*

112 *Id.* at 329.

113 *Id.*

114 See David J. Nicol & Debra Macfarlane-Dick, *Formative assessment and self-regulated learning: a model and seven principles of good feedback practice*, 31 *STUDIES HIGHER EDUC.* 199, 206 (2006) (listing multiple studies showing that students' learning depends on their understanding and ownership of learning goals; "[i]f students do not share (at least in part) their teacher's conceptions of assessment goals (and criteria and standards) then the feedback information they receive is unlikely to 'connect'").

115 See Menendez-Varela & Gregori Giralt, *supra* note 109, at 230 ("[A]n examination of validity should demonstrate that students have successfully overcome cognitive challenges, and that the learning process has been confirmed by various observers—including students.").

116 Sophie M. Sparrow, *Describing the Ball: Improve Teaching by Using Rubrics-Explicit Grading Criteria*, 2004 *MICH. ST. L. REV.* 1, 17 (2004); see also HARLEN, *supra* note 30, at 145 (noting that to limit the fear and emotion aroused by assessment, "we need to be completely open about the need for and purpose of assessment and why it is carried out in particular ways").

B. Use Rubrics as Instructional Tools and to Create a Shared Language of Instruction

Once we have identified which learning goals we will assess and how, rubrics allow us to refine our communication with students. Rubrics describe what students should learn from an assignment and how they will be evaluated;¹¹⁷ they are an excellent way to create a shared understanding¹¹⁸ and to “develop a language” of the course with students.¹¹⁹

Rubrics describe for students the criteria used to assess student performance on an assignment, criteria pulled from the assignment’s learning goals.¹²⁰ Rubrics help us grade more consistently, but they also can be a powerful teaching tool to communicate the specifics of learning goals to students.¹²¹ A rubric supports the consequential validity of an assignment when it informs the dialogue of instruction both in the classroom and in private conferences, and marks student progress toward those learning goals.¹²² A rubric may weaken an assessment’s validity when it has imprecise criteria or descriptions not adapted to the intended audience because when students do not understand the how they will be evaluated, they are less likely to use the assessment to improve on learning goals.¹²³

The discipline of creating accurate, helpful rubrics will focus our assessments.¹²⁴ Creating rubrics requires that we break “learning goals into more specific components that describe what the students have learned”; we describe “what we want students to take away with them, hold on to, and return to in the future.”¹²⁵ We are forced to decide which material is essential and describe a level of performance we expect from our students. This discipline will help us spend our assessment capital knowingly, and communicate that to students.

Rubrics also open communication from the students back to us. One of the troubling features of the Lund University Faculty of Engineering assessment

117 Sparrow, *supra* note 116, at 6.

118 See Menendez-Varela & Gregori Giralt, *supra* note 109, at 228 (“The construction of a shared understanding should begin in the domain of assessment, since this is the major driver for learning—so much so that it concentrates an important part of the student’s attention.”).

119 Sparrow, *supra* note 116, at 21.

120 *Id.* at 7.

121 Menendez-Varela & Gregori-Giralt, *supra* note 109, at 230.

122 See *id.* at 228 (“There is an urgent need to explore rubrics as instructional resources because they are at the service of dialogue, feedback and criticism. When rubrics are conceived as a support for the exchange of expectations regarding learning goals and assignments, and as a support for the judgement of the quality of complex performances, their contribution to validity depends on their degree of integration in a teaching environment that is consistent with the learning goals and student profiles.”).

123 See *id.* at 230.

124 Sparrow, *supra* note 116, at 16.

125 *Id.* at 18.

system is that the teachers did not see when their assessments were hurting students.¹²⁶ Rubrics, when carefully crafted, should communicate what our students are learning, rather than what we believe we are teaching.¹²⁷

C. Ensure the Assessment System Encourages Students' Autonomy

One particularly harmful potential outcome of a tightly orchestrated assessment system is that it may overly dictate student decisions, rather than facilitate student autonomy. A student's feeling of control over his or her learning directly impacts the student's motivation and capacity for learning.¹²⁸ As we add assessments, students should not feel as though they have less control of their learning. Our assessment systems should build students' feelings of autonomy, competence, and relatedness, which are as fundamental to learning as "a plant's need for sunlight, soil, and water."¹²⁹ To facilitate learning, our students should feel they are autonomous, or acting out of their own interests and values; and competent, or working on something they are good at or can become good at; and working relatedly, or connected meaningfully to others in a supportive process.¹³⁰

In their groundbreaking study of student learning, Professors Kennon Sheldon and Larry Krieger found that a law school's autonomy support for students predicted a higher final law school GPA, a better performance on the Multistate Bar Examination, and a higher self-determination to begin a career.¹³¹ Students with autonomy support also demonstrated a higher subjective well-being relative to the baseline.¹³² These outcomes are consistent with wide-ranging studies that have shown that autonomy, competence, and relatedness are "uniquely important in that they have additive effects on a host of positive outcomes."¹³³

According to Professors Sheldon and Krieger, teachers and environments that support autonomy have three defining features: 1) they provide students with as much choice within the constraints of the assignment as possible; 2) they provide students with a meaningful rationale of why the choice may be limited and a sense of the tasks that are required; and 3) they communicate to students that the students' perspectives and concerns are important and

¹²⁶ See Lindberg-Sand & Olsson, *supra* note 72, at 171-72.

¹²⁷ See Sparrow, *supra* note 116, at 18.

¹²⁸ See Fruehwald, *supra* note 96, at 103 (noting that self-determination helps the learning environment and motivation, and adult students in particular want to control their learning).

¹²⁹ See Kennon M. Sheldon & Lawrence S. Krieger, *Understanding the Negative Effects of Legal Education on Law Students: A Longitudinal Test of Self-Determination Theory*, 33 PERSONALITY & SOC. PSYCHOL. BULL. 883, 885 (2007).

¹³⁰ *Id.* at 889.

¹³¹ *Id.* at 894.

¹³² *Id.* at 893.

¹³³ *Id.* at 894.

considered in the decision-making.¹³⁴ As discussed, formative assessments should be grounded in learning goals, they should be incremental, and they should prepare students for summative assessments, but they should not so occupy and control students' learning that students feel as though they are stuck marching to a drumbeat we control or that students become dependent on us to direct their learning activities.¹³⁵

One goal we may have with formative assessments is to teach students to divide larger tasks into smaller ones, and to organize and manage their time wisely. We may use assessment to model and monitor self-regulated learning, which describes a cycle students themselves should implement to guide their learning.¹³⁶ We may, for example, divide up a larger writing project into several smaller assignments or require drafts of a course outline at certain times in a semester. I have found it helpful in my writing classes to require early outlines, but to be flexible with the kinds of outlines I will accept. I encourage student choice by teaching and modeling alternative outlining methods.

If we too tightly control students' process, we may instead stunt their independent progress.¹³⁷ Students need a measure of control to plan their learning, monitor and implement that learning, and evaluate their own learning.¹³⁸ One study on student stress found that time-management skills (setting goals, scheduling, prioritizing, organizing) were effective in lowering college students' stress only when the skills were paired with a feeling of control over their own time.¹³⁹ Students did not feel more in control of their time and less stressed about large projects by the imposition of time-management structures alone.¹⁴⁰ In sum, whenever we structure formative or summative assessments to break up larger tasks for students, we should provide students guidance or a schema, but enough discretion such that they experience control over the pace and order of their learning.

134 *Id.*

135 *See id.* at 884, 889 (defining a controlling environment as one in which a supervisor says, "Tough luck if you don't like it. You have to do it because I say so.>").

136 MICHAEL HUNTER SCHWARTZ, *EXPERT LEARNING FOR LAW STUDENTS* 3 (2d ed. 2008).

137 This controlled approach also runs against a law school's purpose to create practitioners who are metacognitive and self-directed about their own learning. *See generally* Anthony Niedwiecki, *Teaching for Lifelong Learning: Improving the Metacognitive Skills of Law Students through More Effective Formative Assessment Techniques*, 40 *CAP. U. L. REV.* 149 (2012).

138 *See* SCHWARTZ, *supra* note 135, at 3 (describing the three phases of self-regulated learning); *see also id.* at 51-52 & app. A (describing best practices in time management and providing a "self-monitoring log").

139 Sarath A. Nonis et al., *Influence of Perceived Control over Time on College Students' Stress and Stress-Related Outcomes*, 39 *RESEARCH IN HIGHER EDUC.* 587, 589-90 (1998).

140 *Id.*

Likewise, the feedback we give on assessments should support autonomy rather than a parroting back of what students believe we want to hear.¹⁴¹ Paula Manning suggests that autonomy-supportive feedback uses information-rich language that is non-controlling, flexible, and competence affirming.¹⁴² Our detailed rubrics—especially those for formative assessments—should still be open-ended enough to guide students toward solutions, rather than dictating a solution for a student.¹⁴³ Feedback should encourage support and guide students as they become more confident in their independent ability to build a legal proof, to structure an analysis, to counsel a client, and to handle the competing demands of a law practice.¹⁴⁴ Students will not grow into self-directed and determined learners if the rubrics or feedback we use are inflexible and directive.¹⁴⁵ Obviously, we cannot have this role as guide and mentor if students do not sense our concern for their learning and well-being.

Students are likewise less likely to learn from us when they read our feedback and believe they are failures. As discussed *infra*, feedback can be particularly harmful to students with a tendency toward maladaptive responses, but receiving feedback is emotional for most students, even when the feedback is given with advice on how to improve.¹⁴⁶ Just acknowledging this problem may help us stop students from disengaging, but we may also help students by connecting their assessments to the knowledge, skills, and values they will need as practicing attorneys.¹⁴⁷ In other words, ideally our feedback pushes students from focusing on their performance on a specific assessment, and shows them how an investment on improving their performance is possible and will benefit them in their future careers.¹⁴⁸

Taking this sophisticated role in student learning is possible only by becoming more skilled in our assessment practices. As we take more informed control over these assessment practices we will be able to give students more perceived control over their own learning.¹⁴⁹ In a recent study of a school district in rural Pennsylvania working to improve its formative assessment practices, the authors described teachers' evolution moving from consciousness-

141 See Paula J. Manning, *Understanding the Impact of Inadequate Feedback: A Means to Reduce Law Student Psychological Distress, Increase Motivation, and Improve Learning Outcomes*, 43 CUMB. L. REV. 225, 229-31 (2013).

142 *Id.*

143 See *id.* at 233-34.

144 *Id.* at 232-33.

145 *Id.*

146 See Torrance, *supra* note 6, at 334 (noting “assessment is always formative, it will always impact on students and have a central place in what and how students learn, but not necessarily in a positive sense”).

147 See *id.*

148 See *id.*

149 Manning, *supra* note 141, at 232-33.

raising to skill-building and finally to intentional use, where teachers were able to implement formative assessment best practices in their classrooms to intentionally promote student learning.¹⁵⁰ Over the three years of the study, teachers reported students felt more in control of their learning, and students improved in self-efficacy and self-regulation.¹⁵¹ Studies in higher education have likewise shown that students who have become more self-regulated are “more likely to assume control over their learning” and less dependent on external support and direction from the teacher.¹⁵² The more students are positively and authentically engaged— behaviorally, emotionally, and cognitively—the more educationally enabling the classroom environment becomes.¹⁵³

Sheldon and Krieger’s study also demonstrated that students benefit from a broader environment of autonomous support, created by the curriculum and administrative decisions. The study found that the law school with better autonomous support regularly provided teaching skills seminars for its faculty, had more faculty dedicated to practical skills training, and integrated skills and doctrinal professors into the faculty.¹⁵⁴ Students also had a co-curricular requirement that furthered their professional development, and had more practical skills courses to choose from.¹⁵⁵ Sheldon and Krieger concluded that the law school’s entire environment suggested “a stronger orientation toward student interests and priorities,” which led students to perceive more autonomy support and to thrive, as other studies on autonomy support predicted they would.¹⁵⁶

D. Set High Expectations and Display Confidence Students Can Meet Those Expectations

Before the ABA adopted Rule 315 requiring both summative and formative assessments, Professors Carrie Sperling and Susan Shapcott warned that “these changes have missed fundamental concerns that should precede any discussion about feedback: how will students view it, will they learn from it, and could more feedback actually harm students.”¹⁵⁷ According to Professors Sperling and Shapcott, more assessment alone will not help students who

150 Susan Brookhart, Connie Moss & Beverly Long, *Formative Assessment that Empowers*, <http://www.ascd.org/publications/educational-leadership/novo8/vol66/num03/Formative-Assessment-That-Empowers.aspx>.

151 *Id.*

152 Nicol, *supra* note 114, at 205.

153 See Boykin, *supra* note 26, at 508 (citing multiple studies); see also HESS, *supra* note 105, at 261 (an “assessment-centered” environment, in which students have repeated opportunities to practice and receive feedback, is essential to learning”).

154 Sheldon & Krieger, *supra* note 129, at 886.

155 *Id.*

156 *Id.*

157 Carrie Sperling & Susan Shapcott, *Fixing Students’ Fixed Mindsets: Paving the Way for Meaningful Assessment*, 18 LEG. WRITING 39, 41-42 (2012).

are vulnerable to a maladaptive response to feedback.¹⁵⁸ These maladaptive responses include fixed mindset, under which a student believes that feedback is an indicator of his or her set abilities and intelligence,¹⁵⁹ and stereotype threat, which refers to a student's "conscious and subconscious reactions to known negative group stereotypes."¹⁶⁰ Maladaptive responses to feedback limit students' ability to perform, cause students to devalue the skill assessed, and decrease students' enthusiasm and motivation for further effort because of the psychological pressure caused by the fear of confirming negative group stereotypes.¹⁶¹

For students with fixed mindsets, "all assessment is summative rather than formative—a judgment of a fixed trait—not a way to increase intelligence."¹⁶² If law schools fail to address mindset, students with a tendency to a maladaptive response "will continue to respond maladaptively to almost any feedback no matter how carefully it is devised."¹⁶³ Others have shared a concern that early feedback would damage first-year students' enthusiasm for the study of law.¹⁶⁴ Professor Sarah Adams Schoen warned that "the persistent, high-stakes challenge of law school itself may trigger maladaptive responses even when the students are succeeding."¹⁶⁵ A student's success on an assessment does not indicate the assessment lacks validity; but when early feedback on an assessment encourages a maladaptive response, the assessment itself has a low consequential validity because of the damage it does to the student's well-being.

Research on mindset suggests that all students, including those burdened with stereotype threat, may improve their learning by adopting a growth mindset, where students accept feedback as a means to improve.¹⁶⁶ To help

¹⁵⁸ *Id.*

¹⁵⁹ Adams-Schoen, *supra* note 102, at 4.

¹⁶⁰ Russell A. McClain, *Helping Our Students Reach Their Full Potential: The Insidious Consequences of Ignoring Stereotype Threat*, 17 RUTGERS RACE & L. REV. 1, 9 (2016).

¹⁶¹ *See id.* at 5 (noting stereotype threat is a constant limiting force because "of the psychological pressure caused by the fear of confirming negative group stereotypes"); Sperling & Shapcott, *supra* note 157, at 63–64 (warning more feedback for someone with a fixed mindset may increase that student's "vulnerability to depression and anxiety"); Adams-Schoen, *supra* note 102, at 7 ("decades of research in a wide-range of settings and with diverse populations of test subjects strongly suggest that fixed mindset beliefs lead to helpless behavior in the face of perceived obstacles").

¹⁶² Sperling & Shapcott, *supra* note 157, at 58–59, 63 (indicating that their pilot study showed twenty-five percent of incoming law students had a fixed mindset).

¹⁶³ *Id.* at 63.

¹⁶⁴ *See* Emily Zimmerman, *An Interdisciplinary Framework for Understanding and Cultivating Law School Enthusiasm*, 58 DEPAUL L. REV. 851, 898–901 (2009); ROY STUCKEY AND OTHERS, BEST PRACTICES FOR LEGAL EDUCATION (2007).

¹⁶⁵ Adams-Schoen, *supra* note 102, at 17.

¹⁶⁶ *Id.* at 15; *see* Sperling & Shapcott, *supra* note 157, at 47 (noting students with growth mindset

students be more receptive to assessment and feedback, experts recommend assessments have the characteristics already described: that they focus students on learning goals rather than performance goals, encourage student self-efficacy, and give students more autonomy in the learning process.¹⁶⁷ In addition, to specifically target and improve student mindsets, we should teach students about maladaptive responses to feedback,¹⁶⁸ set high expectations for student achievement, provide robust feedback, and communicate confidence that students can meet those expectations.¹⁶⁹

Mindset studies show that this combination of high standards, rigorous feedback, and “express assurances that the students can meet the high standards with effort and persistence” will encourage adaptive responses, including “increased task motivation, trust in the critic, and identification with the skill at issue.”¹⁷⁰ Feedback should avoid generic praise, and should focus on students’ process as much as possible given the assignment.¹⁷¹ Feedback is “rigorous” when it engages with the criteria of the assessment as defined by the learning goals, not when it is simply a reaction to our perception of the student’s effort.¹⁷²

In short, when we address student mindset directly, our assessments are more likely to benefit students’ well-being and increase the likelihood that our assessments will improve student learning in our own courses and in other courses. In contrast, if we simply assess students earlier and more frequently without attention to mindset, we risk triggering maladaptive responses that will broadly harm students’ motivation and their interpretation of our feedback. In that way, maladaptive responses directly impact the validity of our assessments.

E. Regularly Revisit the Assessment System as a Whole, Paying Particular Attention to Students’ Ownership of Their Own Learning

Once we have done what we can to build a coordinated, diverse culture of assessment, we need to ensure the system is responsive by evaluating its impact on students and other faculty members. In the Lund University Faculty of

“approach academic challenges with the goal of increasing their intelligence”).

167 See Sperling & Shapcott, *supra* note 157, at 65–67.

168 See Adams-Schoen, *supra* note 102, at 36–40 (providing a sample five-step lesson plan for an orientation session to encourage long-term mindset changes); Fruehwald, *supra* note 96, at 89 (noting that students will change their thinking and start reacting in new ways just by knowing about the two different mindsets).

169 Adams-Schoen, *supra* note 102, at 41.

170 *Id.* at 32, 41.

171 See *id.* at 41; Sperling & Shapcott, *supra* note 157, at 77–80.

172 See Manning, *supra* note 141, at 244 (recommending genuine constructive criticism, as opposed to criticism grounded in negativity, which can be the source of “anger, fear, embarrassment, or shame”).

Engineering case study, teachers created a coordinated system of assessment but then failed to realize that assessment burdens they created were reducing student learning across the curriculum. Students continued to attend labs and class, and that continued attendance was enough to convince faculty that students were keeping up.¹⁷³

When an assessment fails to improve student learning, we should reevaluate our assessment method, consider the assessment in the context of other pressures on students,¹⁷⁴ and be critical of feedback methods we have used in that or related assignments; but we should also hold students to their own role in the learning process.¹⁷⁵ Sometimes students' habits and expectations are the cause of the failed assessment.¹⁷⁶ What kind of students are we creating: mature, self-directed learners and problem-solvers who can organize and execute learning tasks without our direction, or dependent learners who have yet to understand that learning is something they must do for themselves?¹⁷⁷

If our students—despite our efforts to create a healthy system of assessment—still fail to accept responsibility for their own learning, we may need to alter our classrooms to create a climate that makes clear to students their responsibility for their own learning.¹⁷⁸ The best practices already listed—demonstrating to students our expectations and consequently their work products are tied to discrete learning goals, motivating students with clear rubrics and autonomy, and setting high standards and communicating our belief students can meet them—create this kind of learning environment.¹⁷⁹

Professor Maryellen Weimer recommends two additional steps we can take: First, we must create and students must suffer logical consequences for decisions they make that impact their learning.¹⁸⁰ If students must be prepared to meet the learning goals of a class session, then create an in-class assessment to hold them accountable for this preparation.¹⁸¹ When students in the Lund University Faculty of Engineering faced no clear consequence for missed

173 See Lindberg-Sand & Olsson, *supra* note 72, at 171-72.

174 See Glesner Fines, *supra* note 95, at 97-98 (recommending reviewing “any possible factors that could explain poor performance apart from student learning” to explain deficient student performance on assessments); Zimmerman, *supra* note 58, at 67 (noting that professors could easily discover just by asking whether students are struggling under an unmanageable workload).

175 See Hill, *supra* note 9, at 458.

176 See MARYELLEN WEIMER, *LEARNER-CENTERED TEACHING: FIVE KEY CHANGES TO PRACTICE* 144 (2d ed. 2013).

177 *Id.* at 144-45.

178 *Id.* at 150; see also HESS, *supra* note 105, at 6-7 (describing studies demonstrating that professors' classroom relationship with students influences student learning).

179 WEIMER, *supra* note 176, at 150-58.

180 *Id.*

181 *Id.* at 150-51.

assessments, eventually the assessment system became separated from student learning.¹⁸² Second, we must be consistent with our messages to students about what is expected from them and what they will gain from it.¹⁸³ Students need to view us as reliable and predictable, modeling for them the professional behavior we expect in return.¹⁸⁴ Students should be encouraged to provide the same level of personal investment and introspection into the learning that we do.¹⁸⁵

IV. Conclusion

As law schools expand, improve, and deepen assessment systems to take advantage of the growing evidence that more formative assessment may improve student learning, faculty must continue to measure the validity of their assessments. We cannot assume that more assessment will help students; experts have warned that those students we most want to reach with early assessments are often those students most at risk. Early in their law school career, students may not be prepared to choose how best to spend their time, they may be discouraged by an overwhelming workload, and they may have maladaptive responses to early feedback. Assessment is best added to our curriculum through the best practices developed by our assessment scholars and outside assessment scholars.

But of course nothing—not even scholarly assessment techniques—substitutes for a teacher understanding how his or her students learn and how to intervene in helpful ways.¹⁸⁶ Simply embedding assessments in our curriculum will not improve student learning without our adapting the assessments and our teaching to meet the needs of our course, to accommodate our teaching styles, and to best meet the needs of our students.¹⁸⁷

182 *Cf.* Lindberg-Sand & Olsson, *supra* note 72, at 171-72 (where students were allowed to create these assessment debts with no consequence and the assessment system became separated from student learning.)

183 WEIMER, *supra* note 176, at 152-53.

184 *Id.* at 153; *see also* Hill, *supra* note 9, at 486 (recommending that in conferences we ask students where they lost momentum, whether they devoted the necessary time, whether they are struggling with procrastination, and whether there was an assessment or activity that was more effective with them).

185 *See* Fruehwald, *supra* note 96, at 90-91 (defining motivation as “the personal investment that an individual has in reaching a desired state or outcome”); Paul Black & Dylan Wiliam, *Developing a Theory of Formative Assessment*, 21 *EDUC. ASSESSMENT, EVALUATION & ACCOUNTABILITY* 5, 7 (2009) (noting that because the responsibility for learning rests on both the teacher and the student, each must “do all they can to mitigate the impact of any failures of the other”).

186 *See* Boud, *supra* note 4; Bearman, *supra* note 15, at 50 (encouraging educators to find ways “of developing assessment tasks which take into account their circumstances and pedagogical beliefs and aspirations”).

187 *See* Yue Yin et al., *On the Impact of Formative Assessment, Student Motivation, Achievement, and Conceptual Change*, 21 *APPLIED MEASUREMENT IN EDUC.* 335 (2008) (detailing a study of a science curriculum where formative assessments did not increase student learning because the assessments were added without any adjustments to teaching).