The American Academic Profession

Transformation in Contemporary Higher Education

Edited by

Joseph C. Hermanowicz

The Johns Hopkins University Press
Baltimore

CHAPTER TWO

Focus on the Classroom

Movements to Reform College Teaching and Learning, 1980–2008

Steven Brint

Over the last quarter century, two movements have arisen to encourage a reorientation of the academic profession in the direction of a focus on teaching, rather than research. One has been promoted most actively by foundation-sponsored advocacy organizations, and the other by the federal government and the states. Both of these movements have questioned the priorities of institutions and the preparation and effectiveness of college teachers. They have promoted the idea that college teachers can do a much better job of producing and assessing student learning. The causes of this renewed focus on the classroom are quite similar to those that provoked rethinking of classroom teaching in secondary schools at the turn of the twentieth century: the construction of a mass system, fueled by the incorporation of working-class and immigrant students, in which a majority of students have limited intrinsic interest in learning and in which chronically underfunded schools have limited resources to create powerful learning communities. Expansion, combined with continuing fiscal pressures in the public sector, encouraged concerns about the effectiveness of college teaching, while diversification led to concerns about the possibility of unequal results for women, minorities, and immigrants. While sharing a critical stance toward the current condition of teaching and learning in the academy, the two movements otherwise shared little in common: the one led by liberal philanthropies worked on the improvement of teaching skills, while the state-based movement focused on constructing hard evidence of student learning outcomes.

Higher education policy analyst Peter T. Ewell described the character of the two movements as they emerged in the mid-1980s:

> Two antithetical "ideologies" . . . arose almost simultaneously in higher education discourse. The first came from inside the academy. . . . Its tenets were most clearly stated in an influential national report, *Involvement in Learning* [1984] . . . which argued that breakthrough improvements in undergraduate education could be achieved by establishing high expectations, deploying active and engaging pedagogies, and providing feedback about performance. . . . The second ideology had roots outside the academy based on strong state interest in pursuing [testing-based] educational reform. . . . Its tenets were embodied in a high visibility report by the National Governors Association, *A Time for Results* [1986]. . . . The report argued that colleges and universities should be held accountable for establishing clear standards for performance with respect to student learning and that the results of student assessments should be publicly reported and coupled with consequential actions. (2005, 107)

This chapter shows how the two movements grew out of structural weaknesses in the organization of the academic profession following a period of massive demographic expansion and increased demand on scarce public resources. In the chapter I trace the ideas and projects of the two movements. I also describe the tensions between the major actors in the two worlds of reform—and the commitment of some foundations to work in both worlds. I will emphasize three primary analytical points. The first is simply that large boundary-spanning organizations, always important influences on the research topics of professors, are now working in a serious way to shape the classroom environment. Conventional views about the autonomy of teaching professionals in the organization of classroom life therefore require revision. The second is that variations in favored forms of organization among the competing actors have deeply influenced their ideals, their practices, and their relative levels of success. On one hand, the network-organized, discipline-based, and voluntaristic character of academe has shaped the preferences of the main actors in the teaching reform movement. On the other, the social control interests and metric-driven character of state government have shaped the preferences of the main actors in the outcomes assessment movement. The third is that the strongest force in the environment—stronger thus far than either wealthy philanthropies or powerful state educational bureaucracies—has been the system of mutually reinforcing interests among students, faculty, and administrators that reproduces low achievement standards in many

college classrooms. This system of interests has limited the successes of both movements.

Although the final outcome of the clashes between the two philosophies of reform is as yet unknown, it is clear that one strand of the teaching reform movement gained considerable ground during the period under study. Networks of teaching practitioners have succeeded in disseminating selected principles of what I will call the "new progressivism"—specifically, those principles promoting active learning, civic engagement, and sensitivity to the interests of diverse learners. By contrast, and perhaps surprisingly, the outcomes assessment movement has failed to transform practice, as a result of frequent changes in policy, linked to partisan upheavals, and the capacity of higher education associations and regional accrediting bodies to effectively blunt state preferences for the implementation of standardized performance metrics. Particular disciplines—notably, engineering—have, however, been more completely transformed, as a result of the adoption by professional accrediting agencies of the goals and means of the outcomes assessment movement. The policies adopted by engineering educators could plausibly serve as models for the future.

The primary conclusion of the chapter is that, in spite of tremendous effort over a generation by philanthropies and states, the effectiveness of teaching in American colleges and universities has not changed greatly. What has changed is the rise of progressive education practices. These practices have encouraged greater student engagement, but have shown no strong effects on learning. Indeed, progressive education practices align well with the priorities of student culture, which has been interested in enjoyable activities, but not as interested in demanding requirements and high standards that would help to increase skill development and learning. Another important consequence has been the increased legitimacy attached to the teaching function. This outcome, I will argue, has helped to diminish the status and perhaps also the net social contribution of the profession.

The Contradictions of Postwar Academe

In 1895, William Rainey Harper, of the University of Chicago, was the first American university president to tell newly hired professors that they would be evaluated primarily on the basis of their research contribution. Thus began, in fits and starts, the era of the research-based academic profession. The

emphasis on research intensified in the years following World War II and spread beyond the science and technology fields. By the mid-1960s, the trend toward populating academe with professional researchers was so noticeable that David Riesman and Christopher Jencks coined the term "the academic revolution" to mark what they assumed would be a permanent turning point in the shift of the profession from teaching to research (1968).[1] For research university professors, the requirement to meet the exacting standards of colleagues evaluating articles and books warranted careful training; half-awake, half-interested undergraduates sitting in the back rows of large lecture halls were another matter. In graduate training programs of the period, students were not required to demonstrate skills in pedagogy during their studies for the PhD, nor understanding of the relation between types of pedagogy and subject matter content, nor understanding of the aims or purposes of education. Rather, those who were not fortunate or promising enough to obtain research assistantships were thrown into graduate student-run discussion sections to sink or swim. For most would-be professors, teaching was an amateur activity, performed with limited regard to effectiveness, by people whose real training was for something else entirely.

Many observers within the university welcomed this era of the research-centered professoriate. For Clark Kerr, the new multiversity served the nation by providing greater access, technological progress, and expert advice to every constituency in its state and region. But, Kerr acknowledged, undergraduate teaching suffered: "There seems to be a 'point of no return' after which research, consulting, [and] graduate instruction become so absorbing that faculty efforts can no longer be concentrated on undergraduate instruction as they once were" (1963, 65). Kerr provided no solution to the "cruel paradox that a superior faculty results in an inferior concern for undergraduate teaching," although he hoped that an escape from the paradox could eventually be found (ibid.). More astringent critics, like Jacques Barzun, pointed out the injustice of shortchanging undergraduate students:

> (T)he student . . . is conscious . . . (that his teachers) subject him to cavalier treatment . . . unpunctual, slipshod in marking papers, ill-prepared in lecture, careless about assignments. . . . To put it another way, the student sees and resents the fact that teaching is no longer the central concern of the university. . . . After making all due exceptions (for there are still thousands of devoted teachers and vigilant college heads), the students' complaint is justified.

> The great shift to research after 1945 would alone modify the university atmosphere sufficiently to warrant the impression of neglect, supported as it is by the reality of "publish or perish." (1968, 69)

Although Barzun and others (see, e.g., Schaar and Wolin 1965) expected a student uprising against desultory and negligent undergraduate teaching, these hopes were quickly disappointed. Instead, an ethic of consumerism emerged. In part, this ethic reflected the growth of mass higher education, which brought many more ill-prepared and nonacademically oriented students to campus. The average number of hours spent in class and studying per week dropped from about forty to about twenty-seven in the years between 1961 and 2004. Declines were evident in all institutions, all disciplines, and among all demographic groups (Babcock and Marks, 2010). Moreover, students now had the power, in the form of student evaluations, to register their desires effectively. First introduced in the 1920s, the use of student evaluations of teaching became widespread in the 1970s (Riesman 1980). At large state universities, these forms became the primary method for evaluating performance in the classroom, and they eventually served to encourage faculty to pay attention to the preferences of student consumers for a more entertaining delivery, greater clarity in the structure of lectures, and faculty expressions of kindness and respect. These were improvements over the teaching norms of earlier eras, but student evaluations also encouraged many professors to lower their expectations of student work in the hope of retaining high scores or in response to a declining academic ethos among students (Johnson 2003; Riesman 1980, 249–55).

The contradictions of academic careers were also encouraging renewed attention to teaching among those left out of the "academic revolution."[2] In *The Academic Marketplace*, Theodore Caplow and Reece McGee noted, "For most members of the profession, the real strain in the academic role arises from the fact that they are, in essence, paid to do one job, whereas the worth of their services is evaluated on the basis of how well they do another" (1958, 82). This theme gradually became standard among social scientists writing about higher education (see, e.g., Clark 1987; Ladd, 1979).

In a 1989 national survey of faculty, more than 40 percent of professors strongly agreed that it was difficult to achieve tenure without publishing, up from one-fifth in 1969—and many were not happy about it (Miller et al. 1990). Large majorities at master's- and baccalaureate-granting institutions

said that teaching effectiveness should be the primary criterion in promotion. Nearly as high a proportion of faculty members teaching at smaller doctorate-granting institutions expressed similar sentiments. Only a minority of faculty—those teaching at research universities—could be expected to be rewarded in the labor market for their publications. Others were being required to publish, but with course loads that limited their capacity to do so. Moreover, research funds were not expanding at the same rate as institutional demands for publication. The upshot was that many professors were oriented to teaching and thought they should be granted as much respect as researchers.[3] The academic procession led by the Harvards and Berkeleys was breaking up along institutionally defined lines—dividing those institutions emphasizing research from those emphasizing teaching.

The national context of higher education policy also brought issues concerning the quality of college teaching to the forefront. During the 1980s, the sense that colleges were connected to great national purposes wavered. Policy makers, influenced by the "free market" conservative wing of the Republican Party, began to see higher education as a private consumption good. As demand for credentials grew, some also began to express concerns about educational quality. Some policy makers saw the universities as responding to ill-prepared students with less challenging courses. In many states, political differences between conservative politicians and liberal academics fueled suspicion about the aims and purposes of higher education (Geiger 2004, chap. 2; McLendon, Hearn, and Deaton 2006). In this cauldron of professorial discontent, student consumerism, and Republican Party skepticism, educational quality emerged as a cutting-edge issue.

Teaching Reform Movements

The principal agents of *teaching reform movements* have been foundations and foundation-sponsored advocacy organizations, such as the Association of American Colleges and Universities (AAC&U) and Carnegie Foundation for the Advancement of Teaching (CFAT). These institutions picked up and advanced pedagogical principles in the work of leading educational thinkers of the period. The approach developed by these thinkers, which I will call the "new progressivism," advocated active learning experiences, commitments to diversity and civic engagement, *and* challenging academic standards. However, educators' calls for heightened academic standards proved to be no

match for the consumerist ethos and utilitarianism of college student life. The trajectory of the new progressivism consequently mirrored the pattern of K–12 progressive education in the early twentieth century, when followers of John Dewey, such as William Heard Kilpatrick, de-emphasized Dewey's insistence on rigor and frequent assessment and highlighted the student-centered, active learning, community engagement themes in his work (Cremin 1961).

Teaching Guides

The popularity of guides to good teaching can be seen as one early indicator of change. The National Institute of Education's influential *Involvement in Learning* (1984) signaled both the growing importance of effective teaching and the challenges facing faculty in a system of mass higher education. This document, heavily influenced by the thinking of UCLA higher education professor Alexander W. Astin, advocated movement away from the standard lecture format, so that students could become inquirers—producers, as well as consumers, of knowledge. Following the lead of progressive educators, the report recommended the introduction of "active modes of learning," such as group research projects and classes held in the field, internships and other forms of carefully monitored experiential learning, small discussion groups, in-class presentations and debates, and individual learning projects and supervised independent study. It also advocated timely feedback and more rigorous standards for evaluating student performance (National Institute of Education 1984, 27–28).

Arthur W. Chickering and Zelda Gamson's "Seven Principles for Good Practice in Undergraduate Teaching" represented a similar cast of mind. Their easy-to-remember principles became a touchstone for reformers and formed a basis for subsequent national surveys of student engagement. In a pithy opening sentence, Chickering and Gamson defined the maladies of colleges and universities in an age of mass higher education: "Apathetic students, illiterate graduates, incompetent teaching, impersonal campuses—so rolls the drumfire of criticism" (1987). The seven principles offered something for both progressives (frequent faculty-student contact, collaborative and active learning experiences, and respect for the variety of students' talents and ways of learning) and traditionalists (focus on time spent on task, prompt feedback, and high expectations for performance).

A New Ideology Emerges

Ideologies provide blueprints for action, and by the end of the 1980s organizational changes had created the conditions for an ideological shift—from the research-centered hierarchy of the "academic revolution" to something new reflecting the variety of institutional missions found in U.S. higher education. That new ideology was formulated in Ernest L. Boyer's *Scholarship Reconsidered* (1990). As president of the Carnegie Foundation, Boyer was well positioned to affect the changes in institutional practices he proposed.

Boyer's explicit aim was to install a confederation of interests in the place of academic hierarchy. To do so, he identified four legitimate forms of academic life: the scholarships of discovery, integration, application, and teaching. The use of the venerable term "scholarship" united academe under the idea of studiousness and learning, rather than research and teaching. Indeed, Boyer explicitly hoped to end debates about the relative value of research and teaching. "The most important obligation now confronting the nation's colleges and universities," he wrote, "is to break out of the tired old teaching versus research debate and define, in more creative ways, what it means to be a scholar. It's time [for the profession] to recognize the full range of faculty talent and the great diversity of functions higher education must perform" (1990, xii).[4]

The critical innovation in Boyer's work was the integration of teachers as equal partners in the confederation of scholars. Before Boyer, one rarely thought of teaching as scholarship, only as reflecting knowledge of scholarship. Although the term "scholarship" suggests the possibility of professionalizing the teaching function, for Boyer it remained the province of the inspired amateur, albeit one who thinks deeply about subject matter and reflects often on the effectiveness of his or her practice. Yet the very naming of teaching as a form of scholarship encouraged steps in the direction Boyer himself initially failed to anticipate. More than that, Boyer's essay was a "game-changing" document—the point at which the teaching plebs rose up to challenge the aristocracy of researchers in the name of pluralistic academic democracy. Boyer's book was an academic best seller—the Carnegie Foundation had trouble keeping the book in stock—and Boyer and his colleagues were invited to dozens of campuses and high-profile conferences to discuss the new paradigm he proposed (Glassick, Huber, and Maeroff 1997).

The Rise of Teaching and Learning Centers

Some of the organizational groundwork for teaching reform had already been laid during the period of postwar expansion. The first center for teaching and learning opened at the University of Michigan in 1962, inspired by the work of English and linguistics professors who offered instruction to graduate students on teaching. The faculty senate committee that recommended the establishment of the center observed that opportunities for improving instruction were more important on research university campuses than elsewhere in academe, because most professors were oriented to graduate training and publication.[5]

The form taken by the Michigan center shows many of the characteristics of subsequent faculty-sponsored approaches to improving instruction. It remained voluntary, discipline-based, modular in organization, and reliant on networks of motivated professors to transmit interest and ideas. Michigan resisted "the idea that [faculty and graduate students] are going to be under surveillance" and considered the teaching and learning center's main purpose to be "engendering conversation" (M. Kaplan, personal communication). In 1978, the center began offering orientations to new teaching assistants, and the great majority of teaching assistants at Michigan now receive some common experiences, including discussion of the first days of class; discussion of classroom communication, including teaching in multicultural classrooms; and feedback from analysis of short bits of videotaping. The activities of the center, like those on other campuses, reflected the organization and ethos of academe: the disciplines were preeminent; professors decided how to allocate their time outside of class; and personal interest, rather than university prescription, fueled the enterprise. These strengths of academe, which provide maximum freedom and flexibility for professors to choose their own intellectual and professional paths, were consistent with improved basic training for graduate teaching assistants but arguably led to only a relatively shallow penetration of good practices in the classroom.

By the late 1970s, dozens of universities had opened teaching centers. The Professional and Organizational Development (POD) Network formed in 1975 to provide a professional association for "instructional developers." The largest of the teaching centers, such as those at Ann Arbor, Berkeley, and Austin, provided training to seven to eight hundred new teaching assistants every year. Berkeley required a day-long teaching conference, including five

modular online courses related to pedagogical strategies, ethics, and the educational opportunities and challenges presented by diverse classrooms. In addition, every department offered a seminar on "teaching in the discipline" open to interested graduate students. UT Austin offered minicourses every semester, on a strictly voluntary basis, on topics such as leading discussions and effective lecturing, combined with departmental courses on teaching in the disciplines.

The quality and staffing of teaching centers varied enormously, however. At universities like UC Berkeley with strong demonstrated commitments and relatively stable budgets, well-trained professionals led workshops and provided feedback from videotapes. At budget-strapped campuses, training programs were sometimes led by mentor teaching assistants who were themselves just learning their craft. Nor did all campuses mandate teacher training orientations. In 2001, one-third of research universities said they required no mandatory orientations for teaching assistants (Reinvention Center 2002). Moreover, seminars on teaching in the disciplines were rarely required, according to respondents; only 10 percent of responding research universities in 2001 said they required such seminars (ibid.).[6]

Institutionalizing the New Progressivism: AAC&U and NSSE

The Association of American Colleges and Universities (AAC&U), which defined itself as the only major national organization focusing on liberal and general education, added the theme of diversity to the new progressivism and became one of the most important agents of change in the undergraduate curriculum. During the 1980s and 1990s, the vision of AAC&U focused on reshaping the liberal arts to bring diversity within the compass of the fundamental commitments of liberal education.[7] In the early 1990s, AAC&U effectively advocated the addition of courses on gender, diversity, and non-Western cultures to the general education curriculum (see, e.g., Cornwell and Stoddard 1999; Musil 1992). The organization saw itself as a "leading edge of change" whose goal was to "amplify what [it] sees in the field" (D. Humphreys, personal communication).

This work culminated in the American Commitments initiative (1993–2001), funded by the Ford Foundation, the Hewlett Foundation, and the National Endowment for the Humanities. The connection between diversity and democracy provided a signal theme for this work. AAC&U drew on familiar images of pluralism, but with a new twist: "Higher education," it wrote, "can

nurture Americans' commitment and capacity to create a society in which democratic aspirations become democratic justice. Diversity proves a means of forging deeper civic unity" (Beckham 2000, 2). This conceptual link between diversity and democracy brought diversity thoroughly into the mainstream of liberal education, while updating the Dewey tradition to incorporate the race- and gender-conscious movements on campus.

The AAC&U developed powerful organizational tools to realize its vision. These included the formation of a national panel, composed of prestigious figures in academe, modeled on the blue-ribbon commissions that had long been used by the federal government as a means for focusing support for policy initiatives. They also included Diversity Leadership Institutes, held during the year and, more intensely, during summers for teams from twenty to thirty member institutions. These institutes disseminated best practices for reforming general education as a vehicle for teaching about diversity and for promoting "global social awareness." They also included community seminars to "discuss and re-imagine what it means to be a citizen in a multiracial society" (www.aacu.org). AAC&U was one of the first to find effective use of the Web for creating compendia of campus practices and resources to promote diversity and for highlighting successful efforts to implement changes in organizational practices. Its flagship magazine, *Liberal Education*, highlighted diversity initiatives on member campuses and the connection between diversity and democracy at the heart of the American Commitments initiative. The association claimed that 160 campuses were involved in at least one of the face-to-face programs and that 100 institutions undertook efforts to rethink curriculum and to provide opportunities for students to consider "critical questions about American pluralism." The association itself grew from six hundred to eight hundred members during the period of the diversity and democracy initiative.

AAC&U's efforts to update progressive education ideals for the twenty-first century took a new turn in the late 1990s as it confronted the challenges of the state-based accountability movement. Its new initiative took up the fundamental issue, what should be the characteristics of a liberally educated person in the twenty-first century? In its projects, AAC&U promoted a new vision of liberal education combining traditional aims with progressive ideals and a new conception of twenty-first-century skills. The program built cleverly on the dynamic new force of perceived employer dissatisfaction with the

qualifications of college-educated labor,[8] and it mobilized support for alternatives to standardized testing of student learning outcomes.

Funded by four foundations (the Carnegie Corporation of New York, the Charles Engelhard Foundation, the Pew Foundation, and the John Templeton Foundation), together with the federal Fund for the Improvement of Postsecondary Education, AAC&U offered what amounted to a tripartite solution to the re-creation of liberal education, blending exposure to the traditional core fields of knowledge (natural and social science, humanities, and arts), cross-curricular work on cognitive and expressive skills (analytical and critical reasoning, written and oral communications, quantitative and information skills), and commitment to the values of educational progressivism (intercultural understanding, personal development, civic and social engagement, and integrative and collaborative learning).

Like the American Commitments initiative before it, the Liberal Education and America's Progress (LEAP) initiative drew on a familiar set of opinion formation mechanisms mastered by powerful Washington lobbies: reports of national panels of distinguished academics and business leaders, Web site resources extolling the values of the new policy agenda, community forums to discuss the new vision, and magazine articles focusing on the implementation of campus reforms reflecting the new vision. Both AAC&U magazines, *Liberal Education* and *Peer Review*, took up the cause. These magazines were sent to five administrators on member campuses, broadening exposure within the communities of change agents on the campuses. By 2009, organizational membership had grown to 1,200 institutions, each one sponsoring five campus representatives; these 6,000 campus representatives connected to AAC&U through member institutions and periodical subscriptions constituted a core of reform-minded activists spread through academe.

The new vision had much to do with countering the growing threat of state regulation of the college classroom through standardized testing. Together with its new vision of the essential skills and values for the twenty-first century, the initiative brought a new approach to assessment to the fore. This new approach focused not on standardized testing, along the lines of K–12 accountability, but instead on electronic portfolios and senior capstone courses. As the association's main report on accountability and assessment stated, "Capstone courses and portfolios provide promising anchors for a meaningful approach to educational accountability" (AAC&U 2004, 8). This work left out the details about how

students' course work could be fairly sampled or assessed for improvement over the college career or examined for evidence of proficiency in specified outcome areas. These matters were to be left to the colleges, to develop in ways that fit local conditions, rather than by independent third parties.

The National Survey of Student Engagement (NSSE) represented another powerful force in the institutionalization of the new progressivism. Led by George D. Kuh, a professor of higher education at Indiana University, NSSE was launched with Pew Foundation funding in 2000. NSSE built on decades of research by Kuh and his colleague Robert Pace on the College Survey of Educational Quality (CSEQ) (Kuh 2009). This work closely paralleled the precepts of *Involvement in Learning*. Conceived in part as an alternative to the resources- and reputation-based college rankings of *U.S. News and World Report*, NSSE intended to measure more accurately the actual quality of undergraduate students' educational experiences. The five NSSE benchmarks, each addressed through scaling related questions, probed levels of student-faculty contact, active and collaborative learning, academic challenges, educational enrichment activities, and institutional climates conducive to learning.

In its inaugural year, NSSE was administered at more than 270 institutions; this number grew to more than 600 annually by the end of the decade (www.nsse.iub.edu). Institutions were soon comparing their engagement scores on the five key dimensions to national norms and norms for institutions of their type. NSSE generated an impressive number of reports detailing the distribution and consequences of engagement experiences, and it also championed case analyses of institutions that showed exceptional effectiveness in the production of engaged learning environments (Kuh et al. 2005). However, NSSE measured engagement, not learning,[9] and although many college educators assumed that higher levels of engagement should register more or less directly in improved learning outcomes, empirical efforts to demonstrate this proposition were disappointing. Student scores on NSSE scales were, for example, only weakly associated with scores on the Collegiate Learning Assessment (CLA), and most factors failed to reach statistical significance once students' prior academic records (grade point average and SAT scores) were controlled (Carini, Kuh, and Klein 2006).[10] Other studies showed that high grades were more common in humanities and social science courses, in which the culture of engagement emphasized participation, interaction, and active learning experiences, and were less common in the natural sciences and engineering, where engagement typically meant long hours of

study, with groups of peers, to master demanding quantitative material (Brint, Cantwell, and Hanneman 2008).

Promoting Teaching for Understanding: CFAT

The forces of the new progressivism had impressive organizational tools under their command and a relatively easy-to-implement checklist of reforms to attach to existing curricula. The same could not be said of the more ambitious and less completely realized project of the Carnegie Foundation for the Advancement of Teaching under Ernest Boyer's successor, Lee S. Shulman. Under Shulman's leadership, the Carnegie Foundation embarked on a program to redefine and realize Boyer's vision of a scholarship of teaching. These efforts eventually steered the foundation away from the tenets of the new progressivism to a deeper inquiry into aims and methods of undergraduate teaching. Shulman's approach came to share only part of the faith of the new progressivism in the power of student engagement. Engagement, he wrote, "is not enough." "Understanding is not independent [of engagement] but is an additional standard" (Shulman 1989).

For Shulman, all good teaching was built, in the first instance, on subject matter mastery. Shulman emphasized, in addition, "pedagogical content knowledge"—the special materials and methods tied to knowledge making in the disciplines, such as work with primary textual materials in history, surveys and ethnography in sociology, and diagnostic clinical rounds in medicine. Based on this knowledge and these disciplinary resources, teaching and learning could be conceived as an interactive process of bringing "something inside" of the teacher out in a methodical and powerful way, as well as bringing "something outside" of the student, the lesson, into strong relief in students' consciousness. In all good teaching, methods of expression and bases of apprehension and understanding were consequently closely linked (Hutchings and Shulman 1999).

Shulman and his colleagues emphasized that the first obligation of the teacher is to determine what students know and can do, as well as their interests and passions. Working from these bases, Shulman and his associates (Huber and Hutchings 2005) advocated that teachers create "cognitive apprenticeships" in which students are asked to make their mental processes accessible to their fellow students and teachers and to work toward more expert understandings of course materials. Through a process of "uncoverage," teachers were encouraged to focus their first lessons on ideas and concepts that were both

difficult to grasp and fundamental to subsequent learning in the class. Teachers made their own thinking accessible to students by explicating the "intermediate processes" of understanding—the understandings that are employed habitually by expert learners but are often hidden in the process of instruction. These could include, for example, explicit discussions of the flow of an argument or text, the translation of terms no longer in wide use, or a detailed, step-by-step interpretation of the architecture of a statistical table. Other techniques for making knowledge accessible included slowing down students' reading to elicit students' descriptions of their thinking about passages in text; administering oral rather than written midterms; employing structured online discussions to create learning communities oriented to key issues and ideas in a course; and posting examples of beginning, intermediate, and advanced understandings of texts with detailed explications of the major differences between these levels of mastery. Similar pedagogies were developed for mathematics—for example, in James Sandefur's "think alouds," in which math students were asked to describe, step by step, how they were thinking about a problem as they worked through its solution.

Shulman argued that students should demonstrate competence by performing skills in front of their teachers and classmates, rather than by passively absorbing information. For Shulman, the pathologies of learning—amnesia (forgetting what was just learned), fantasia (misperceiving the lesson to reinforce existing knowledge), and inertia (inability to use knowledge in new contexts)—were ultimately issues of ownership. Understanding implied ownership and ownership typically required performance (see also Shulman 1997).

The institution of the Carnegie Academy for the Scholarship of Teaching and Learning (CASTL) was the first of Shulman's organizational vehicles. CASTL was based on the idea that reform began in small groups, rather than as a broad ideology. It sought not to transform but to create strong emotional loyalties among those who self-selected as reformers. The Pew Foundation provided a $5 million grant to Carnegie to inaugurate CASTL.[11] Pew funds provided support for a summer academy located at the foundation where successful applicants, approximately fifteen a summer, met together to discuss and develop the ideas from their proposals for improvements in teaching and learning. The projects ranged widely, but most sought to understand the learning process or to develop conditions under which broader and deeper learning could occur in classroom settings. They included, for example, a project by the English teacher Mariolina Salvatori to develop the idea of "difficulty papers," in which students identify and begin to hypothesize the reasons for a possible difficulty they experience reading a poem, play, or essay. Another project, by the psychologist Jose Feito, mapped the conditions for more broadly distributed learning in seminar settings, including ways of helping students take responsibility for "owning" the learning process, building appreciation of multiple perspectives, creating expectations for the contributions of all members, and creating a space in which students could safely acknowledge their lack of understanding.

Growing out of the Carnegie program, Scholarship of Teaching and Learning (SoTL) colloquia sprouted up on hundreds of college and university campuses during this period. These colloquia took up visually effective presentation of lessons, new ways to assess student learning, uses of technology to improve pedagogy, the impact of learning communities, and many other topics consistent with the Carnegie agenda. On most campuses, SoTL sought to foster discussion and incremental change based on emulation of appealing approaches to the challenges of teaching. Rooted in the precepts of the constructivist pedagogy, the philosophy was not an industrial search for better systems, but rather an apprenticeship system for craftsmen, based on sharing the distinctive visions of master teachers.

Other Shulman-inspired projects led to the creation of Web sites intended to spread pedagogical practices consistent with the "teaching for understanding" approach. Georgetown professor Randy Bass's *Visible Knowledge Project* Web site (1999) was the most important for advancing and codifying ideas about pedagogies of understanding. Bass obtained a five-year, $2.6 million grant from Atlantic Philanthropies "to improve the quality of college and university teaching through a focus on student learning and faculty development in technology-enhanced environments." His Web site spotlighted techniques for slowing down and deepening knowledge transmission, for building on core ideas and concepts, and for making teachers' intermediate processes and performance standards visible to students, while revealing students' prior understandings and making their difficulties in understanding course materials visible to teachers.

Shulman's interest in updating the shop talk of teachers included advocacy of electronic "teaching commons" where proven ideas could be "documented, shared and built upon" and thereby gain wider currency (Shulman 1993). In 1995, University of Nebraska professor Dan Bernstein launched the Web site *Peer Review of Teaching* to realize Shulman's goal of making teaching

"community property." Peer review of teaching, as developed by Bernstein, began with the exchange of three memoranda between colleagues. These memoranda discussed the objectives of the courses, the instructional design for the course, and the quality and breadth of student understanding demonstrated in the course. Based on these memoranda, Bernstein's Web site allowed college teachers to document what they did in their classes through electronic course portfolios. The portfolios included graded examples of student work, representing a range of how well students had achieved course goals. They also included a section identifying the next steps in the development of the course. As in other Carnegie-inspired efforts, the Web site touched the work lives of only relatively small numbers of devoted practitioners. They did, however, help to launch the leading alternative that emerged during the period to standardized testing of student learning outcomes—electronic teaching portfolios, which included detailed analysis of course-level assessment data.

The organizational apparatus Carnegie used to spread these ideas showed neither the panache of the AAC&U campaigns nor the reach of NSSE. Instead, an artisanal model, built on networks of sympathetic practitioners, prevailed. This approach generated fresh insights about teaching and learning—insights with the potential to create more effective college teachers. But its insistence on "scaling down" through small-scale actions of unusually committed practitioners was destined to create islands of improved practice in a sea of relative indifference. According to Huber and Hutchings, "the key is not the scale and scope but the care and thoughtfulness of the work, its capacity to change thought and practices, its generosity, even, perhaps, its power to surprise and delight" (2005, 30).[12]

Carnegie itself changed dramatically with the selection of Anthony Bryk in 2007 to replace the retiring Shulman. Bryk launched an effort to "scale up" R&D in education through well-supported industrial-style prototyping and mass diffusion, beginning with a project to turn around the low success rates of community college students in remedial mathematics. This represented a sharp departure for a foundation modeled under Shulman as a think tank for teaching craftsmen. Russell Edgerton, who did so much as a program officer at the Pew Foundation to promote Shulman's agenda, concluded ruefully that more than two decades of reform activity sponsored by liberal philanthropies like Pew had resulted in "neither professional nor institutional transformation" (R. Edgerton, personal communication).

Outcomes Assessment Movements

I now turn to the other movement to reform college teaching and learning, that led by the states. Outcomes assessment can be defined as a response of state governments and regional accrediting bodies to the perception that colleges and universities have not done enough to ensure that students are learning course materials and essential academic competencies. Where the teaching reform movement took root in foundation-supported advocacy organizations, the *outcomes assessment movement* was promoted primarily by the states and the federal government.[13] Following the K–12 reform model, state officials have sought to investigate these issues using relatively low-cost, quantitative measures. Policy think tanks, such as the National Center for Public Policy and Higher Education and the National Center for Higher Education Management Systems (NCHEMS), played important roles articulating and promoting the objectives of the assessment movement. Interinstitutional higher education associations, such as the National Association of State Universities and Land Grant Colleges (NASULGC), later renamed the Association of Public and Landgrant Universities (APLU), and the American Association of State Colleges and Universities (AASCU), have attempted to mediate between universities and the states, as have the regional and disciplinary accrediting bodies. Both the higher education associations and the regional accrediting agencies followed the goals of the assessment movement by insisting on evidence of student learning outcomes. The regional accrediting agencies allowed institutions and disciplines to define their own measures of student learning outcomes, while the higher education associations developed a voluntary system of accountability that allows participating institutions to choose from three authorized assessment instruments to test "core academic skills."

Fledgling efforts to encourage institutional assessment of learning outcomes began in the 1970s. The Educational Testing Service (ETS) fielded the first open-response test of core skills, Academic Competencies in General Education, tested at 140 institutions, but later abandoned as a result of the tendency of institutions to magnify small pretest/posttest differences and the test's unreliability in the midranges of scoring (Adelman 2007). By the mid-1970s, twenty states had introduced minimal competency testing for graduating seniors, mirroring popular high school exit exams (Gilman 1978). Calls for action continued in the early 1980s. *A Nation at Risk* (1981) documented the shortcomings of U.S. primary and secondary education in the face of increasing

competition from East Asia. Only four years later, *A Time for Results* (1985) stressed the same fears about the competency of U.S. college graduates and the same looming threat of Asian competition. It noted that U.S. higher education had set a new standard for access, but observed that "access without quality is a cruel deception." In the document, a subcommittee of governors, led by John Ashcroft of Missouri and including future president Bill Clinton of Arkansas, questioned common assumptions about higher education: "Learning is assumed to take place as long as students take courses, accumulate [credit] hours and progress satisfactorily toward a degree." But, the subcommittee observed, "tests of elementary and high school teachers show that the BA is not a guarantee of even basic literacy, let alone competence." The report also cited, with little documentation, "substantial levels of dissatisfaction" among employers about the skills of college graduates. The report advocated systematic programs using multiple measures to assess undergraduate student learning, and it cited with approval institutions like Alverno College that had pioneered systematic assessment in the 1970s. It also applauded the Southern Accreditation Commission for being the first of the regional accrediting bodies to require an assessment component for reaccreditation.

Performance Funding: The First Wave

Beginning in the 1980s, states began to demand that universities account in detail for the ways they were spending their money, the amount professors were teaching, and, to a lesser degree, how much students were learning. A study team led by political scientist Michael McLendon reported in the 1980s that state financial resources became conditioned upon institutional performance in specified areas. These often included student retention and graduation rates, student scores on licensing examinations, job placement rates, faculty research productivity, and measures of undergraduate access and campus diversity (McLendon, Hearn, and Deaton 2006). Between 1979 and 2007, twenty-five states enacted performance funding, but ten of those states dropped it during the period (Burke and Minassians 2003; Dougherty and Reid 2007). Performance funding proved costly to implement, susceptible to institutional manipulation of performance measures, and subject to reversal under new administrations, or when unstable state finances caused deep cuts in regular higher education funding (Burke and Serban 1998; Dougherty and Natow 2009; Shulock and Moore 2002; Zumeta 2001).

Nevertheless, new demands for accountability, including direct assessment of student learning, slowly gained ground during this period. A 1983 report of the Education Commission of the States showed that two-thirds of states had initiated some form of required student assessment. However, many states used minimal competency measures at graduation, or even more indirect measures, such as graduation rates and pass rates on professional licensing examinations. Although assessment of student learning was in the air, few knew how to test directly for student learning outcomes in a cost-effective, relatively unobtrusive way. Regional accrediting agencies, like the North Central Association, began requiring institutions to plan for ways to directly assess evidence of student academic achievement, and state higher education policy think tanks, such as ECS and NCHEMS, issued statements of support for the endeavor. The large testing companies, ACT and ETS, also geared up for the new era by introducing or revamping multiple-choice tests, the Collegiate Assessment of Academic Proficiency (CAAP) and the Measure of Academic Proficiency and Progress (MAPP), respectively, that institutions could administer to their freshmen and seniors to determine the institution's "value added" to student academic competencies.

Pressure on state budgets contributed to this sharper focus on the college classroom. In the 1990s, state appropriations for higher education declined for the first time in real terms. Although funding recovered in the later 1990s, the recovery was slow and shallow, and state appropriations fell steeply in real terms with every new recession. In the context of limited and unstable revenue bases and stiff competition for public dollars, some state governments began to demand performance assessments in return for funding commitments (Alexander 2000). These state actors wanted to know whether they were receiving value from their investments in higher education.

As critics of government waste continued to score points, Democratic Party centrists argued that government could become much more efficient by monitoring the performance of its functional units closely, with an eye for creative ways to meet consumer service goals. David Osborne and Ted Gaebler's *Reinventing Government* (1992) became a popular guidebook for state reformers; its animating ideas were endorsed by Vice President Al Gore and others who were interested in defusing long-standing concerns about the wastefulness of government spending. Broader trends in the appropriation of powers to regulate professional work were also at play. State officials emphasized the tendency of unregulated professionals to feather their own nests and

prescribe overly expensive treatments. In many cases, doubts about the ectiveness of professional practices, combined with the increasing cost of oviding services, led to third-party regulation, greatly reducing the auton- ny of professionals.[14] State dissatisfaction with control of work by the occu- ıtional community was, in this respect, another central context out of which igher education accountability movements grew.

A Bandwagon Forms

To the extent that an analogue exists in the outcomes assessment move- nent to Ernest Boyer's *Scholarship Reconsidered*, it was produced by two state ollege professors in California, Robert Barr and John Tagg, in a widely cited 995 article from *Change* magazine. In this article, Barr and Tagg sought to shift thinking in academe from an "instruction paradigm" to a "learning para- digm": "In the briefest form, the paradigm that has governed our colleges is this: A college is an institution that exists to provide instruction. Subtly but profoundly we are shifting to a new paradigm: A college is an institution that exists to produce learning. This shift changes everything" (1995, 1). The new paradigm fostered a change in focus from a pedagogy based on instructors' expectations about what students should learn to one emphasizing what stu- dents actually do learn. The article advocated sophisticated assessments grounded in "minds-on" problem solving, while it explicitly supported exter- nal evaluations of learning. Barr and Tagg looked forward optimistically: few external evaluations of learning had, to this point, focused on sophisticated assessments of learning; instead, nearly all sought inexpensive ways to assess student learning through multiple-choice instruments.

The idea of a shift to a "learning paradigm" resonated strongly among state educational bureaucrats and in the world of higher education policy analysts. By 2001, ten states, concentrated in the South and Midwest, had experi- mented with or adopted standardized testing of student learning outcomes (Ewell 2001b). The idea of demonstrating how much institutions added to learning was gaining widespread appeal. Advocates of this "value-added" ap- proach argued that this could be done by controlling for predicted gains based on student "input" characteristics at college entry, such as their social back- grounds and SAT scores, and then attributing residual gains to institutions. However, few in the policy community agreed on what types of learning should be measured or how it should be demonstrated. Some argued for

discipline-specific knowledge, others more general cognitive skills (such as ana- lytical thinking and writing), and still others wanted to focus on work-related skills. Some advocated multiple-choice tests for their cost-effectiveness, but others concluded that higher-level cognitive skills could not be demonstrated in this context and required the completion of more complex, "real-world" tasks.

In spite of disagreements about what should be tested, the learning out- comes movement gained traction as higher education leaders in the national associations and regional accrediting bodies concluded that they could no lon- ger ignore state pressures to "show results." The list of supporters for increased accountability included many of the foundations that were simultaneously supporting projects to reform teaching. Already active in promoting "assess- ment forums" at the annual meetings of the American Association for Higher Education, the Pew and Danforth Foundations provided grants to regional accrediting agencies in 1999 to work on criteria for collection of data on stu- dent learning outcomes.

Over the next five years, a chorus of influential voices called for measure- ment of student learning outcomes and created demonstration projects to show how this measurement could be done. In 2000, the National Center for Public Policy and Higher Education, funded by several major foundations and led by the former governor of North Carolina and educational reformer James B. Hunt, began to publish report cards about state higher education performance, including "incomplete" grades for all states on student learn- ing. In the same year, the Accreditation Board for Engineering and Tech- nology (ABET), the accrediting agency for engineering schools, began its *Engineering Criteria 2000* policy, requiring outcomes measures and plans for continuous improvement based on results of outcomes assessments. In 2002, the Pew Trusts provided funding to two leaders of the assessment movement, Margaret Miller and Peter Ewell, to demonstrate the possibility of measuring college learning in six states for future incorporation into the National Center for Public Policy and Higher Education's "Measuring Up" reports. In 2003, the Carnegie Corporation of New York and the Teagle Foun- dation sponsored the development of a new type of test of core academic skills, the Collegiate Learning Assessment, based on the use of document libraries to solve "real-world" problems. In the same year, the national council of regional and disciplinary accrediting agencies, the Council for Higher Edu- cation Accreditation (CHEA), announced a policy of "mutual responsibility"

between institutions and regional accrediting agencies for demonstrating student learning outcomes.

An opinion survey published by ETS, also in 2003, discovered evidence of public concerns about educational quality, stronger among political conservatives and high school–educated people. Primed by questions linking costs to quality assurance, a majority surveyed by ETS agreed that colleges should provide evidence that they were producing the learning results they promised, if they were going to continue to raise costs (ETS 2003). In 2004, the Business–Higher Education Forum argued for the first time in favor of assessments of student learning outcomes. Also in 2004, the State Higher Education Executive Officers launched a National Commission on Accountability in Higher Education, chaired by former secretary of education Richard Riley and former Oklahoma governor Frank Keating, both Republicans. The report they produced in 2005 concluded that most state systems "do not meet their intended purpose to improve and to provide evidence of student learning" and endorsed collection of data on student learning outcomes (National Commission on Accountability in Higher Education 2005). In the same year Miller and Ewell published their six-state report showing that states could demonstrate student learning outcomes through a variety of measures, such as proficiency benchmarks modeled on the K–12 National Assessment of Educational Progress (Ewell and Miller 2005).

The Spellings Commission and the VSA

Buoyed by this swelling interest in higher education accountability, the Bush administration turned its attention from K–12 reform to higher education. Secretary of Education Margaret Spellings appointed a Commission on the Future of Higher Education, chaired by Texas businessman Charles Miller, to recommend reforms in higher education accountability. In 2004 and 2005, the commission issued a number of preliminary reports critical of higher education's commitment to transparency, cost containment, and, most important, demonstration of results for student learning. In 2006, the commission issued its final report, *A Test of Leadership*, which was highly critical of the performance of America's colleges and universities. The report dismissed previous efforts to bring accountability for student learning outcomes. "Despite increased attention to student learning results by colleges and universities and accreditation agencies, parents and students have no solid evidence, comparable across institutions, of how much students learn in colleges or whether they learn more at one college than another. Similarly, policymakers need more comprehensive data to help them decide whether the national investment in higher education is paying off and how tax payer dollars could be used more effectively" (Commission on the Future of Higher Education 2006, 14). The commission advocated measuring student achievement on a value-added basis that took into account students' previous achievements when assessing outcomes. It stated that this evidence should be made available to consumers and policy makers in an accessible, understandable way, and it encouraged the implementation of "meaningful" interstate comparison of student learning in all states (ibid., 4).[15]

The specter of high-stakes testing haunted many in academe, who argued that such tests would yield little of value for students studying such a wide variety of disciplines (see, e.g., Chatman 2007; Hawthorne 2008). The only way to test learning would be discipline by discipline, these educators argued, and this seemed an impossible task given the limited resources of colleges and universities and the limited capacity of state educational bureaucrats to grade such a wide variety of tests. An article by the assessment expert Trudy Banta summarized the experience of educators who had attempted to implement standardized tests of general intellectual skills, such as interpretation, critical analysis, and writing. Banta argued that such instruments primarily test entering ability, are not content neutral and therefore privilege students specializing in some disciplines more than others, contain questions and problems that do not match the learning experiences of all students at any given institution, and measure at best 30 percent of the knowledge and skills that faculty want students to develop. She also raised doubts, based on her own research, about the reliability of gain scores at the individual level, the extent to which students take such tests seriously, and the dangers posed by high-stakes testing for narrowing the higher education curriculum to focus on the skills and content emphasized in the tests (Banta 2007).

Leaders of the testing movement countered that tests of general skills were an important, if not the only important, measure of student achievement in college. Instead of relying on one test, they argued, multiple forms of assessment would be necessary—some to assess general skills, others to assess disciplinary knowledge, and still others to assess the "soft skills" required in leadership positions (see, e.g., Ewell 2004; Shulenburger 2008). Institutions

could be responsible for these assessments, provided that they took their responsibilities seriously.

Following publication of the Spellings Commission report, attention in Washington shifted to the struggle over the reauthorization of the Higher Education Act of 1966, which had been languishing in Congress since 2003. The Bush administration, which had already placed several accountability-minded trustees on the national Council for Higher Education Accreditation, proposed that the federal government take a larger role in quality assurance. Some influential senators, including a leading Democrat, Edward M. Kennedy, argued for bringing higher education into an accountability structure parallel to that of No Child Left Behind (NCLB). As in the case of NCLB, Kennedy wanted to focus on the education of minority and first-generation students by tying increased federal spending to increased federal responsibility for quality assurance. Following extensive lobbying by the higher education associations, Senator Lamar Alexander, a former secretary of education, was convinced to allow the existing system of voluntary accreditation to continue and to bar the federal government from prescribing standards that these agencies were required to use in assessing institutional effectiveness. But, in exchange for his support, Alexander insisted that higher education institutions themselves take on the responsibility to measure student learning outcomes in a serious way.

The reauthorization passed without an enhanced federal role. Alexander's intervention led to the creation of the Voluntary System of Accountability (VSA), organized, with support from the Lumina Foundation, by two of the leading higher education associations, APLU and AASCU. The creators of VSA were very clear about wanting to avoid an NCLB-type system in which important subjects might be driven out of the curriculum. They were also very clear about the need for a voluntary system until such time as the construct validity of existing assessments could be definitively established. Finally, they were aware of the pressure they were under from state and federal education officials, who believed that the time had long passed for higher education institutions to take accountability seriously. As David Shulenburger, the vice president for academic affairs of APLU, put it: "Our detractors allege that we are unproductive, wasteful, and that our students benefit less than we have claimed. . . . If it accomplishes nothing else, generating and publishing transparent, comparable, and meaningful data will serve to diminish the volume of those who believe we are hiding something" (Shulenburger 2008, 21–22). VSA set as an explicit goal the development of a system of accountability that would

"facilitate comparisons of learning outcomes among institutions of higher education."

Testing companies were quick to sense the opportunity to expand their higher education markets. ETS sponsored a national advisory panel to discuss the virtues and defects of existing instruments. ETS issued two reports on "creating a culture of evidence" (Dwyer, Millet, and Payne 2006; Millett et al. 2007). The first of the reports was influenced by the debate surrounding the Spellings Commission and the reauthorization of the Higher Education Act, the second by the triumph of the VSA approach. On the basis of the second report, the creators of VSA chose three tests as acceptable measures of institutional "value-added" to core academic skills: ETS's own Measurement of Academic Proficiency and Progress, ACT's Collegiate Assessment of Academic Proficiency, and the Council for Aid to Education's Collegiate Learning Assessment.

Of these three, the CLA elicited the most interest among policy makers and others who wanted to compare institutions. Like MAPP and CAAP, the CLA tested capacities for analysis and synthesis, not simple recall, but it tested these capacities using document libraries and real-life scenarios, rather than the true/false and multiple-choice format of more conventional instruments. Specifically, the CLA asked students to complete a performance task and two analytical writing tasks. Each performance task had its own document library that included a range of sources, such as letters, memoranda, research reports, newspaper articles, maps, and photographs. The performance task required students to answer open-ended questions about "a hypothetical but realistic situation." One sample question asked students to evaluate whether available data tend to support or refute claims about weaknesses in the construction of the wing of an airplane that a fictitious company was planning to purchase for its sales force. The analytical writing tasks required students to make and critique arguments. One sample question asked students to make an argument that responded to the following claim: "There is no such thing as 'truth' in the media. The one true thing about the information media is that it exists only to entertain." Another asked students to evaluate whether fast-food restaurants contribute to childhood obesity based on a report about a research study.

As a measure of higher-level general skills learning outcomes, the CLA had clear strengths in comparison to multiple-choice tests, but it was also not without some weaknesses when used as a measure of institutional "value

ed." The creators of the CLA claimed that value-added information could obtained with samples as small as one hundred. Given the small sample s permitted by VSA, it was impossible to know whether differences among itutions were due to the composition of student samples by major fields of dy or other student characteristics, differences in motivating incentives, or e institutional differences in educational effectiveness. Although the cre- rs of CLA controlled for incoming SAT scores, they did not require controls the disciplinary composition of samples. Samples composed mainly of nmunications majors, for example, would likely perform rather differently m samples composed mainly of engineering majors. The test therefore at- buted "value-added" to institutions in some cases in which changes might more accurately attributed to disciplinary or other more specific educa- nal experiences.[16] Moreover, the CLA did not regularly report total error or nfidence intervals (see Braun 2009). In state systems, less prestigious insti- tions tended to show greater gains than more prestigious institutions. The eators of CLA denied that ceiling effects could be a factor in these results, t the fact remained that most entering freshmen at state flagship universi- s scored high on the test before any institutional effects came into play. udents at less prestigious branches of the university scored low and had ore ground to gain.[17] Many institutions put off implementation of tests of re academic skills prescribed by the VSA. Of the more than three hundred stitutions participating in VSA as of fall 2009, less than one-third had re- orted results of "core academic skills" using one of the three authorized testing struments. Of those institutions reporting results, the expected two-thirds eported results within a standard deviation for institutions with similar student cademic ability profiles, but, oddly, among the remaining institutions three imes as many reported results "above" (one standard deviation above the redicted mean) or "well above" (two standard deviations above the predicted nean) as reported results "below" or "well below" expected. Only 5 of 104 eporting institutions said that they were performing below expected levels.

An Incremental Approach: CHEA and the Regional Accrediting Bodies

The six regional accrediting agencies are organized and directed by aca- demics (or former academics) as quality assurance agencies. The system was developed as an explicit alternative to state regulation of higher education. Although regional accrediting agencies are independent of the states, they are subject to state recognition, which has proven to be an important

lever. In 1989, federal regulations first required accrediting organizations to examine student learning outcomes as a condition of recognition. The efforts of the regional accrediting agencies to implement review rubrics and to train peer reviewers were aided by funds from the Pew Foundation. By the mid- 1990s all six of the regional accrediting agencies had policies in place requir- ing institutions to demonstrate not only that they were tracking conventional measures of student success, such as four- and six-year graduation rates, but also that they had mechanisms in place to achieve established goals for stu- dent learning. In 1998, Congress formalized this commitment by making stu- dent achievement the first of nine areas in which the regional accrediting agencies were required to have standards.

Even as they followed federal directives, regional accrediting bodies buff- ered colleges and universities from state pressures to introduce standardized testing. Some of the "regionals" have allowed institutions to take responsibility for assessing and achieving a unique set of learning outcomes that institutions have established for themselves. Others have named a core set of learning out- comes that ought to be examined by all institutions. These typically encom- passed, at a minimum, critical and analytical thinking, written expression, and quantitative reasoning. Institutions and departments have been granted considerable autonomy so long as they provide evidence that they are establish- ing learning objectives and developing ways to assess and report the achieve- ment of these objectives. This permitted a variety of assessment approaches, ranging from the presentation of portfolios of student work to requirements for integrative research papers in senior capstone courses. Others built in learning objectives to required courses and required samples of work from these courses or adopted exit examinations as a way of determining whether learning objec- tives had been met. Although the regional accrediting bodies developed elabo- rate procedures to ensure that institutions did more than pay lip service to their demands for evidence of student learning, their requirements were neverthe- less often treated as an encumbrance requiring the appearance of compliance without deeper commitments to the goals of evaluating student learning in a broader way than class grades allowed. The limited resources and experience of accrediting agencies also encouraged institutional autonomy; most, if not all, lacked experience in evaluating evidence of student learning or the quali- fications to establish clear standards by which to do so (Ewell 2001a).

Even so, the regionals created much more attention to student learning out- comes than had existed before. In 2009, the National Institute for Learning

itcomes Assessment (NILOA), housed at the University of Illinois, fielded a idy of the incorporation of assessment instruments. The study was funded the Carnegie Corporation, the Lumina Foundation, and the Teagle Foundation. Officials at half of U.S. two- and four-year institutions responded to e survey, and the vast majority (92%) said that they were engaged in stitution-level assessments of student learning. Most said they were using rvey instruments like NSSE, but 39 percent said they were also using standardized tests of general knowledge and skill like the CLA. At the program vel, four of five respondents said they were assessing student learning outmes in at least one program, and here portfolios dominated. Most said that creditation was the primary driver of their interest in assessment (Kuh and enberry 2009).

Engineering, with its competency-based outlook and favorable attitude toard operational planning and evaluation, provided the most ambitious echanism for transforming undergraduate education through reforms deeloped by its disciplinary accrediting board, ABET. ABET's *Engineering Criteria 2000* (*EC 2000*) offered both a more prescriptive orientation to expected utcomes of the undergraduate curriculum and stronger mechanisms for lanning and demonstrating achievement of these outcomes. Specifically, *EC 000* required detailed published educational objectives, a process in which bjectives were determined and evaluated, a curriculum that ensured achievenent of these objectives, and a system for using results of assessments for ontinuous improvement of the effectiveness of the program. In addition, it stablished specific outcome criteria that all engineering graduates were, in heory, required to demonstrate. These included the ability to apply knowledge of mathematics, science, and engineering; the ability to design and conduct experiments, as well as to analyze and interpret data; the ability to design a system, component, or process to meet desired goals; and the ability to identify, formulate, and solve engineering problems. The criteria also included social and communication skills, such as the ability to function on multidisciplinary teams, to understand professional and ethical responsibility, to communicate effectively, and to demonstrate knowledge of contemporary issues (ABET 2000).

ABET's focus on learning led to changes in instruction. In an evaluation of *EC 2000*, between one-half and two-thirds of faculty surveyed reported that they had increased their use of active learning methods, such as group work, design projects, case studies, and application exercises, to meet learning objectives. In this study, a comparison of 1994 and 2004 engineering graduates showed small but significant self-reported gains in technical abilities, such as the application of mathematics and science to engineering problems. Students also self-reported more sizable increases in social areas specified by *EC 2000*: ability to work in teams, understanding of professional ethics, understanding of contemporary issues, and global cultural awareness (Lattuca, Terenzini, and Volkwein 2006). Outside of engineering, the controls imposed by accrediting agencies were still relatively weak by the end of the decade, but they were slowly changing the way institutions thought about the outcomes of higher education. Most institutions were engaged in assessing their contributions to student learning. Undergraduate program reviews had been institutionalized across the country, and although these varied dramatically in quality, they provided regular feedback to departments based on external, third-party review. Departments have been required to think, sometimes seriously, about what they expect students to gain from their programs and to provide at least skeletal evidence that these objectives were being met. The "audit culture" (Tuchman 2009) spread with each new reaccreditation.

Consequences of the Two Reform Movements

What in the end have the two movements for reform of college teaching and learning produced? The answer to this question depends on whether we look at their practical consequences or their consequences for the legitimacy of teaching work as the central identity and activity of academic professionals.

Practical Consequences of the Reform Movements

It is safe to say that preparation for classroom teaching improved during the period, thanks to the diffusion of basic training for graduate teaching assistants through the auspices of teaching centers. When William Cummings and Martin Finkelstein surveyed U.S. faculty in 1992, they found that only 30 percent of respondents said they had any training for teaching before they took their first jobs (Cummings and Finkelstein 2007). The proportion of graduate students receiving basic training for teaching has now more than doubled in recent cohorts (Boice 1992; Golde and Dore 2001; Reinvention Center 2002).[18]

Classroom practices also changed dramatically in the direction advocated by the new progressives, even as part of their message was lost. Here the best

data come from the Higher Education Research Institute's (HERI) triannual studies of the American faculty. From the late 1980s through the mid-2000s, extensive lecturing showed a marked decline as a teaching method, even in public research universities, and cooperative (small group) learning opportunities showed a corresponding increase. Full-time college faculty increasingly said they were bringing their students into field settings, asking them to demonstrate their knowledge in front of class through oral presentations, relying on reflective writing and journaling, using real-life problems to illustrate lessons, and putting student-centered inquiry, rather than recitation of facts and concepts, at the center of their teaching work (Astin, Dey, and Korn 1991; DeAngelo et al. 2007; Dey et al. 1993; Lindholm et al. 2002, 2005; Sax et al. 1996, 1999). These changes have gone together with an expanded conception of the goals of undergraduate education. Consistent with principles of the new progressivism, the *American College Faculty* studies also show sharp increases in the centrality of social goals: reaching out to surrounding communities through community-based research, teaching appreciation of multicultural diversity, and interest in using undergraduate education as a vehicle for promoting social change. Just as the twentieth-century progressives socialized their ideals of citizenship through the schools, so too do college faculty now overwhelmingly endorse the goals of diversity and community engagement.

These preferences were evident in all segments of American four-year colleges and universities, as much in private colleges as in public universities. The main proponents of these changes have been younger and female faculty members (DeAngelo et al. 2007, 5, 9, 11), suggesting that the trends are likely to continue as older faculty retire and college teaching faculties become increasingly populated by women and those brought up in the norms of the new progressivism.

Active learning experiences reflect a time-honored way to engage the interests of students—particularly less academically oriented students—and are, in this sense, responsive to the changing demography of undergraduate student bodies. The changing demography of the professoriate provides complementary support. At the same time, the checklist character of progressive education has likely also mattered in its widespread adoption. Professors can ask themselves and mentally check off whether they have added hands-on learning experiences, collaborative learning projects, and readings that are responsive to diverse learners.

Active learning pedagogies have apparently not led to great change in student learning, however, at least insofar as this can be measured by students' performance on the CLA. Looking at a sample of 2,400 students who took the CLA at the beginning of their freshman and middle of their sophomore years, sociologists Richard Arum and Josipa Roksa (2011) found that students had improved their critical thinking, complex reasoning, and writings skills, as measured by the CLA performance task, by only .18 standard deviations, or an average 7 percent gain. Forty-five percent of students showed no statistically significant change in their CLA scores. Arum and Roksa concluded that students' completion of three semesters of college had made a "barely noticeable" impact on the higher-level cognitive skills tested by CLA. Follow-up work on the same sample found that, even after four years of college, 36 percent of students showed no statistically significant change in their CLA scores.

Trend data from NSSE provide clues about why this may be so. These data show that many active and collaborative learning activities have grown more popular over time, while challenging requirements, such as the amount of time students spend studying per week and the number of twenty-page papers they write, have remained static or fallen (NSSE 2000, 2008). In the 2008 NSSE report, nearly two-thirds of seniors in NSSE sample institutions said they studied fifteen or fewer hours per week, and half said they had never written a paper of twenty pages or longer (NSSE 2008). In both cases, challenging requirements were less common in 2008 than those found eight years earlier.

The acquiescence of faculty to the preferences of student consumers explains many of these trends. Students have effectively resisted professorial demands for higher levels of effort by simply refusing to engage their studies at a deep level. Ethnographic studies indicate that students have relied on posted lecture notes, the prevalence of relatively easy courses to fill out their schedules, and teachers' openness to negotiations concerning work demands and grades (see, e.g., Grigsby 2009; Moffatt 1989; Nathan 2005). Arum and Roksa report that more than 90 percent of students say they have talked to a professor about grades, but only one-quarter say they have talked to a professor about ideas presented in class. A majority of the 2,400 college students in the Arum and Roksa study said they had not taken a course during the previous term that required a total of twenty pages of written work, and 25 percent said they had not taken a course that required even forty pages of reading per

week. Arum and Roksa conclude that students learn little because they do not study much and little is demanded of them. The college experience, they argue, is perceived by many students as at its core a social experience, rather than an academic experience.

The current low expectations system of undergraduate education does not accurately describe the practices common at some liberal arts colleges or in some of the more demanding disciplines, such as engineering, math, and physics. But it does accurately describe the system of undergraduate education in most institutions and in a majority of non-STEM fields. The system exists because it serves the interests of all major actors who are in daily contact with the classroom. The majority of students see college as a period of fun, friendship, and personal development before they begin adult life. They hope their investments in college-level training will pay off in the labor market, but many assume that credentials will add value, not what they have learned in college. While faculty members are interested in making their classes lively and interesting, they also want to preserve time for research, correspondence, committee work, and other socio-professional activities. Often the livelihoods of adjunct faculty depend on high student evaluations. Administrators at nonselective institutions have been more interested in reaching enrollment targets and raising retention and graduation rates than in encouraging challenging course work or requiring students to demonstrate cognitive growth (Arum and Roksa, 2011; Bok 2006; Brint 2009).

The states have proven to be strong advocates of assessing student learning outcomes, but weak implementers. Early efforts to assess student learning outcomes focused not on direct evidence, but rather on such indirect measures as retention and graduation rates, pass rates on state licensing examinations, and student satisfaction surveys. Today, the states have been persuaded to defer to the regional and professional accrediting associations to provide quality assurance and to the VSA to experiment with the construct validity of several tests of general intellectual skills and to use these tests to monitor the "value-added" of institutions.

Neither the regional accrediting bodies nor the VSA have as yet transformed the college classroom by demanding evidence of student learning outcomes. Richer discussions are underway now about learning objectives, but the regional accrediting agencies have, for the most part, allowed institutions to choose their own

of regional accreditation, dependent on the state but responsive to the voluntarism, decentralization, and discipline-centered character of academic life. Similarly, the learning outcomes component of VSA has been slow to get off the ground. Its champions have wanted to allow for debate and discussion, and they have purposefully insisted on voluntary participation. But institutions have also dragged their heels when asked to provide evidence that could jeopardize their claims to excellence. VSA has also been plagued by doubts about the validity of value-added tests as compared to criterion-referenced tests of competence. Thus, while national and trans-institutional actors have succeeded in shaping the environment of discussion, their efforts have met both passive and active resistance whenever they have attempted to prescribe tough standards for the assessment of student learning outcomes.

Political considerations appear to have had an important influence on the preservation of teaching autonomy in higher education. These political considerations include the ability of higher education advocates to exploit doubts about the effectiveness of K–12 reform, partisan turnover in the governing coalitions of the states, and, in particular, the capacity so far of higher education associations and regional accrediting bodies to assure key legislators that they would implement accountability measures responsive to public interest in quality assurance. Finally, most states do not currently have the resources to fund third-party implementation and scoring of tests like the CLA.

Thus, the most obvious consequences of two decades of reform have been the diffusion of active learning pedagogies and surface-level adoption of relatively weak accountability measures. These will clearly not be enough to change the social relations of learning currently prevailing in most college classrooms. Instead, improvement will require the establishment of higher expectations and more challenging course requirements. They will also require wider penetration of the practices of teaching for understanding developed by Lee Shulman and others. A taste for confrontation with student culture will be essential for college teachers to make progress in improving students' academic skills and stimulating their interests in the life of the mind, as will rigorous assessment of the success of their efforts.

Consequences of the Legitimation of Teaching Identities

Even as course requirements leveled off or fell, the academic profession's self-concept was effectively altered by ideologies that placed teachers, rather

and accountability

nents have had perhaps their greatest success in raising the legitimacy of ing as an object of concern and as a central identity for academics. In the recent national survey of postsecondary faculty, more than three-quarters fied teaching as the most important activity in their professional lives ıster and Finkelstein 2008, 87). The faculty as a whole reported that 60 nt of its work time was spent on average in teaching-related activities, as ared to 15 percent on research (ibid., 88). Only the natural and social sci- and engineering showed any reapportionment of effort in the direction of rch (ibid., 91). In addition, institutions more often required evidence of hing excellence" in applications for positions; such evidence was required percent of advertisements placed in the *Chronicle of Higher Education* (Mei- and Kaplan 2008). These requirements grew at all types of institutions, ıding research universities, and particularly in the arts and humanities.

he establishment of teaching as an accepted core identity for professors ed the problem of status inconsistency (prestige for research, but require- its mainly for teaching) first identified by the sociologists Theodore Ca- v and Reece McGee. But it also augured an era in which the academic ession devolved both in its aspirations and in its accomplishments. Ernest er wished to maintain scholarship at the center of the profession. Yet the erican College Faculty surveys suggest that the centrality of scholarly con- utions has itself slowly eroded in the face of the participatory practices and emosynary goals of professors. Among full-time faculty in public doctoral- nting universities, interest in becoming an authority in one's field declined 10 percent between 1989 and 2004, before increasing a bit in 2007. Interest obtaining recognition from colleagues for scholarly achievements showed imilar rate of decline. American college faculty outside of private universi- s were more likely to say in 2007 that helping others was a more important al than becoming an authority in one's field or obtaining recognition from lleagues for scholarly contributions (DeAngelo et al. 2007).

These data suggest that support for teaching did not preserve scholarship the unifying feature of the academic profession, as promised in *Scholarship* reconsidered, but rather that college teaching was transformed from more of scholarly profession into more of a helping profession. This transformation as aided not only by the decoupling of the teaching-centered academy from ıe research-centered academy but by the success of a modern version of educa- onal progressivism that catered to the interests of students in undemanding ıasses while reducing requirements for student performance.

At the same time that they begin to confront the consumerist and utilitar- ian norms of student culture, American academics may soon find it necessary to recreate the research-centered hierarchy of the post–World War II era. After decades of U.S. dominance, in recent years European scholars have taken over the lead in scientific publication. During the 1990s, the EU15 over- came the United States as the world's most productive region of scientific work. Where U.S. scientists produced nearly 40 percent of papers in the early 1970s, their share was down to one-quarter by the mid-2000s (National Sci- ence Foundation 2007; see also Galvez et al. 2000). Although the United States remains far ahead in articles with the highest citation rates, this gap is also closing (Horta and Veloso 2007). Other countries have improved their infrastructures for scientific production and the quality of their graduate pro- grams. According to a 2007 survey, U.S. professors reported less time spent on research than professors in a number of countries, including Canada, Japan, Korea, Hong Kong, and China (Cummings and Finkelstein, cited in Jaschik 2009).[19] Moreover, three out of five U.S. professors characterized themselves as leaning toward teaching, rather than research, as their primary involve- ment, as compared to 30 to 40 percent of professors surveyed in five other developed countries (Canada, Hong Kong, Japan, Korea, and the United King- dom). When weighing their involvement in teaching against research, the pro- file of U.S. professors resembled that of Brazilian and Mexican academics more than that of professors in developed countries (W. Cummings, personal communication).

These data are disturbing if one believes that the transmittal of the skills and practices of research and scholarship are at the center of the social con- tribution that university professors can make. Of course, the top research uni- versities and liberal arts colleges will maintain a primary focus on the values of scholarship and the powers of mind that scholarship develops. But this is a narrow circle of institutions, and many faculty members at less prestigious institutions, empowered by Ernest Boyer and his followers, have been led to challenge its influence as elitist and remote from the everyday problems of students. As the academic profession has divided, the more numerous teach- ing group has begun to develop its own nonscholarly norms of practice. In non-STEM fields, faculty members have gravitated to active learning experi- ences and social service goals. Boyer expected pluralism to strengthen the usefulness and unity of the profession. But one might well ask, in the wake of the unintended consequences produced by *Scholarship Reconsidered*, whether

a strong academic profession can be one whose sense of itself is focused more on progressive pedagogies than the scholarly disciplines. College professors can and should continue to improve their teaching practices. But one can only hope that professors will resist calls for reform that undermine the ultimate source of their profession's strength, its commitment to the standards of scholarly research.

NOTES

I would like to thank Scott Patrick Murphy for research assistance on this project. I would like to thank Richard Arum, William Cummings, Kevin J. Dougherty, Russell Edgerton, Martin Finkelstein, Mary Taylor Huber, Matthew Kaplan, Wendy Katkin, George Kuh, Mindy Marks, and Ernest Pascarella for bibliographic suggestions and access to unpublished data that helped to improve the quality of this chapter. I would like to thank Philip Babcock, Trudy Banta, Henry Braun, Steve Chatman, Kevin J. Dougherty, Judith Eaton, Russell Edgerton, Peter T. Ewell, David Fairris, Joseph Hermanowicz, Mary Huber, Jal Mehta, Margaret Miller, Gary Rhoades, Josipa Roksa, Carol Schneider, Jack Schuster, David Shulenberger, and Lee S. Shulman for conversations and correspondence that helped to challenge and advance my thinking about issues of teaching and learning in U.S. higher education.

1. Christopher Jencks and David Riesman described "the machinery" for producing researchers in *The Academic Revolution*: "[The top universities] have long been remarkably similar in what they encourage and value. They turn out Ph.D.s who . . . mostly have quite similar ideas about what their discipline covers, how it should be taught, and how its frontier should be advanced. . . . These men were not only like-minded at the outset, but they have established machinery for remaining like-minded. National and regional meetings for each academic discipline and subdiscipline are now annual affairs, national journals publish work in every specialized subject, and an informal national system of job placement and replacement has come into existence. The result is that large numbers of Ph.D.s now regard themselves almost as independent professionals like doctors or lawyers, responsible primarily to themselves and their colleagues rather than their employers, and committed to the advancement of knowledge" (1968, 13–14).

In the postwar era, Jencks and Riesman contended, college and university presidents ceded control to these professional men: "The typical president's greatest ambition for the future is usually to 'strengthen' his institution, and operationally this . . . turns out to mean assembling scholars of even greater competence and reputation than are now present" (1968, 17).

2. Market conditions also contributed to the renewed interest in the craft of teaching. The market for full-time faculty appointments turned markedly more competitive in the tighter years following the great enrollment expansion of the 1960s and early 1970s. While the number of positions for new faculty remained roughly constant as a result of retirements and separations, new entrants faced markedly different circumstances for two reasons: the number of newly minted doctorates was growing much faster than positions for them—with larger cohorts of nearly 15,000 a year by 1997 (Schuster and Finkelstein 2008, 164)—and many more college teaching positions were being created off tenure track (ibid., 194). Between the 1970s and 1997, cohorts of new PhDs grew from fewer than 30,000 a year to more than 42,000 a year. Whereas most hiring had been on tenure track in the 1970s and 1980s, most new faculty members were being hired off the tenure track by the early 1990s (ibid.). The competition led graduate students to consider how best to give themselves an edge in the competition for faculty jobs. For students seeking jobs in research universities, this meant increased efforts to expand professional networks and to publish during graduate school. But some graduate students realized that evidence of teaching ability could constitute a plus factor that might tip appointment committees in their favor.

3. By the mid-1980s, clear signs were emerging of erosion in the "academic revolution" ideal of a research-centered profession. A study of department chairs by Burke (1988) revealed that research qualifications and research potential remained the most significant criteria used in hiring assistant professors, but that teaching ability had become an important part of the equation everywhere. Baccalaureate- and master's-granting institutions, in particular, were looking more and more at teaching as the primary criterion for hiring, even as research universities remained focused on publication and research potential.

4. The scholarship of discovery—or basic research—was, in Boyer's framework, the distinctive activity of professors in the arts and sciences of leading research universities, and particularly those working in the natural science disciplines. The scholarship of application—or applied research—was the distinctive activity of professors in professional schools at research and doctoral-granting institutions. It is the effort to apply knowledge to the solution of problems—"whether in medical diagnosis, serving clients in psychotherapy, shaping public policy, creating an architectural design, or working with the public schools" (23). The scholarship of integration—or synthetic interpretation—was the distinctive activity of humanistic scholars working in liberal arts colleges and research universities. Such scholars "give meaning to isolated facts, putting them in perspective." This was not, he cautioned, the work of the gentleman scholar or dilettante, but rather "serious, disciplined work that seeks to interpret, draw together, and bring new insight to bear on original research" (19).

5. Another initiative began at Harvard and a few other leading universities during the same period based on elective courses in teaching problems and methods offered to graduate students. At its peak, the Harvard course enrolled one in eighteen graduate students (Barzun 1968, 35).

6. The proportion of both universities requiring teaching assistant orientations and those requiring seminars on teaching in the disciplines dropped by the end of the decade (W. Katkin, personal communication).

7. As early as 1969, it had issued a statement crediting minorities for "giving a fresh and compelling impetus to the movement for restoring relevance to academic programs" (AAC 1969). Its studies on the "chilly climate" for women in college classrooms (Hall and Sandler 1982) received national attention in the 1980s.

8. Previous polls had shown employers to be relatively happy with higher education (see Lusterman 1977; Zemsky and Iannozzi 1998) and more interested in the development of social presentation skills and conformity than in the development of cognitive skills (see, e.g., Lesgold, Feuer, and Black 1997; Lusterman 1977; Squires 1979). AAC&U embraced the cognitive skills agenda and publicized its own poll of business executives, conducted by the Democratic pollster Peter D. Hart, showing that CEOs whose businesses employed high proportions of college graduates were in accord with the AAC&U agenda (Peter D. Hart and Associates 2006). If the poll results were unbiased—far from a certainty—businessmen and educators were, perhaps for the first time, developing a community of interest in the outcomes of higher education. Undoubtedly, workplace concerns helped to facilitate such rapport as existed, including a concordance of interest in collaborative and small group learning and intercultural understanding in increasingly diverse workplaces.

9. NSSE included student self-reports of learning gains in several skills areas. Self-reports show modest correlations with objective tests of learning gains and cannot be taken at face value as evidence of student learning (see, e.g., Bowman, forthcoming).

10. A similar study with more elaborate controls on students' prior achievements also yielded modest or insignificant relationships between NSSE benchmarks and cognitive growth on the Collegiate Assessment of Academic Performance (Pascarella, Seifert, and Blaich 2009).

11. Russell Edgerton, who moved from AAHE to the Pew Foundation in 1997, played an instrumental role in the institutionalization of the Carnegie reforms. Edgerton had "discovered" Shulman in national conference presentations in the 1980s and had become a devotee of Shulman's ideas for improving teaching and learning in academe. At Pew, Edgerton worked closely with colleagues at Carnegie throughout the decade of Shulman's presidency.

12. The total number of CASTL scholars topped out at fewer than one hundred. SoTL colloquia emerged on campuses throughout the country, but they attracted only a minority of motivated teachers to their events. Even at such a highly engaged campus as Indiana University, only about one-quarter of tenured and tenure-track faculty had participated in a SoTL event by 2002, and fewer than sixty people attended these events, on average, on a campus of more than two thousand faculty members. The Visible Knowledge Project ran out of funds in 2005, after a decade of pioneering work. Peer Review of Teaching remained operational, but attracted a dwindling number of new portfolios after Pew funding ended. Carnegie's Knowledge Media Lab closed its electronic doors in September 2009, although its course portfolio software remained retrievable.

13. Outcomes assessment should be distinguished from the broader movement to increase accountability in higher education. Accountability has been linked to such performance indicators as graduation and job placement rates, as well as learning outcomes. Performance funding, a popular approach to provide incentives for improved institutional performance, is an outgrowth of the broader accountability movement (see, e.g., Burke 2005; Dougherty and Natow 2009).

14. In response to the breakdown of the ideal type of professional autonomy, understood as occupational control of work, sociologists have proposed a variety of alternatives to preserve professionalism or to reconfigure it for the contemporary world. These include blueprints for bolstering the ideological and ethical underpinnings of the professions (Freidson 2005), suggestions that professionals be able to demonstrate empirically that occupational control of work leads to better results for their clients than market or state-bureaucratic control (Brint 2006), and more comprehensive reworking of the professional model as part of a cooperating joint enterprise involving contributions from a variety of related occupations (Adler, Kwon, and Hechscher 2008).

15. Disappointing results from the National Assessment of Adult Literacy (NAAL) were one cornerstone of the commission's case for improved measurement and monitoring of student learning outcomes. NAAL data seemed to show that only 30% of college graduates could accurately interpret two competing editorials or make accurate inferences from a graph relating age, exercise, and blood pressure. Later administrations of the test to samples made up exclusively of recent college graduates showed no declines in literacy. The National Research Council concluded that the test as constructed could not detect who was proficient in literacy skills (National Research Council 2005).

16. The CLA and similar assessment instruments focus on important cognitive abilities related to analysis, synthesis, and evaluation. This strength of the CLA was not well aligned with two of the most important traditional aims of higher education: to provide general education in basic fields of knowledge and advanced training in a specialized discipline. In the past, every "high-stakes" test has brought a focus on the skills and content it privileges and only on those skills and contents. Indeed, the designers of the CLA acknowledged that they would be happy if colleges and universities taught to their test (see, e.g., Shavelson 2007). Some observers consequently argued that widespread adoption of the CLA or similar instruments would lead to the reconstitution of college classrooms around document-based performance tasks and tasks that involve making or breaking an argument (see, e.g., Brint 2008), at best an incomplete approach to undergraduate education.

17. It is perhaps not surprising under the circumstances that results for students at the El Paso and Permian Basin branches of the University of Texas showed higher

n expected gains on the CLA, while those at UT Austin did not, or those at the versity of North Carolina–Wilmington showed higher than expected gains while se at the University of North Carolina–Chapel Hill did not (see www.collegepor lt.org/#).

18. The adequacy of preparation and, especially, pedagogical mentoring during duate school remained open to doubt. Most PhD-granting institutions provided y limited incentives to improve teaching practice, beyond one-day orientation rkshops. Semester courses on teaching in the disciplines remained uncommon. ey were mandated at no more than 10% of universities (Reinvention Center 02), and few teaching assistants were closely monitored for their work in the assroom. In an online survey, Golde and Dore (2001) found that fewer than 40% 32,600 responding doctoral students reported that teaching assistants in their ograms were adequately supervised to improve their teaching skills. Depart- ent chairs continued to indicate that incoming faculty would benefit from addi- onal training in teaching (Benassi, O'Brien, and Seidel 1998; Meizlish and Kaplan 008).

19. The triumph of the new progressivism may reflect a broader change in so- ial values. In 1999, nearly half (47%) of Americans saw science and technology as he country's greatest achievement. That proportion slipped to just over one- quarter (27%) in 2009. By contrast, civil rights and equal rights were seen as Amer- ca's greatest achievement by 17% of the population in 2009, up from 5% a decade efore (Pew Center for People and the Press 2009). While the election of Barack Obama as president may help to explain these polling results, the results also sug- gest that the eroding prestige of science—and, with science, the research base of the university—could become a real concern. Increased public concern for equal opportunity has gone hand in hand with an increased emphasis on pedagogies of engagement and social goals in the university. These changes may or may not be compatible with the continuing centrality of cognitive rigor as the core value of the university.

REFERENCES

Accreditation Board for Engineering and Technology (ABET). 2000. *Criteria for accrediting engineering programs*. Baltimore: ABET, Inc.

Adelman, Clifford. 2007. Death to value-added. www.insidehighered.com/views/2007/01/26/banta (accessed October 30, 2007).

Adler, Paul S., Seok-Woo Kwon, and Charles Heckscher. 2008. Professional work: The emergence of collaborative community. *Oganizational Science* 19:359–76.

Alexander, F. King. 2000. The changing face of accountability. *Journal of Higher Education* 71:411–31.

Arum, Richard, and Josipa Roksa. 2011. *Academically adrift: Limited learning on college campuses*. Chicago: University of Chicago Press.

Association of American Colleges (AAC). 1969. *Racial problems and academic programs*. Washington, DC: AAC.

Association of American Colleges and Universities (AAC&U). 2004. *Our students' best work: A framework for accountability worthy of our mission*. Washington, DC: AAC&U.

Astin, Alexander W., Eric L. Dey, and William S. Korn. 1991. *The American college teacher: National norms for 1989–90 HERI Faculty Survey*. Los Angeles: Higher Education Research Institute, Graduate School of Education, UCLA.

Babcock, Philip S., and Mindy Marks. 2010. The falling time cost of college: Evidence from a half century of time use data. *NBER Working Paper No. 15954*. Cambridge, MA: National Bureau of Economic Research.

Banta, Trudy. 2007. A warning on measuring learning outcomes. *Inside Higher Education* (January 26). www.insidehighered.com/views/2007/01/26/banta (accessed October 12, 2008).

Barr, Robert B., and John Tagg. 1995. From teaching to learning: A new paradigm for undergraduate education. *Change* 27 (November/December): 12–25.

Barzun, Jacques. 1968. *The American university: How it runs, where it is going*. Chicago: University of Chicago Press.

Bass, Randall. 1999. The scholarship of teaching: What's the problem? www.doit.gmu.edu/Archives/Feb98/randybass.htm (accessed July 3, 2009).

Beckham, Edgar F. 2000. *Diversity, democracy, and higher education*. Washington, DC: Association of American Colleges and Universities.

Benassi, Victor A., E. J. O'Brien, and Lee F. Seidel. 1998. *The 1998 national survey of new faculty hiring practices*. Paper presented at the Lilly Conference on College University Teaching. Boston.

Boice, Robert. 1992. *The new faculty member: Supporting and fostering professional development*. San Francisco: Jossey-Bass.

Bok, Derek. 2006. *Our underachieving colleges*. Princeton: Princeton University Press.

Bowman, Nicholas A. Forthcoming. Can first-year college students provide accurate self-reports about their learning and development? *American Educational Research Journal*.

Boyer, Ernest L. 1990. *Scholarship reconsidered: Priorities of the professoriate*. Princeton: Carnegie Foundation for the Advancement of Teaching.

Braun, Henry. 2009. Finding the value in value-added. Paper presented at the 3rd SERU Research Symposium. University of California–Berkeley (May).

Brint, Steven. 2006. Saving the soul of professionalism: Eliot Freidson's institutional ethics and the defense of professional autonomy. *Knowledge, Work, and Society* 4:99–123.

———. 2008. The Spellings Commission and the case for professionalizing college teaching. *Academe* 94 (May-June): 21–24.

———. 2009. Student culture in an age of mass consumerist higher education. Unpublished lecture given at Teachers College, Columbia University (March).

Brint, Steven, Allison M. Cantwell, and Robert A. Hanneman. 2008. The two cultures of undergraduate academic engagement. *Research in Higher Education* 49:383–402.

Burke, Dolores L. 1988. *A new academic marketplace.* Westport, CT: Greenwood Press.

Burke, Joseph C., ed. 2005. *Achieving accountability in higher education: Balancing public, academic and market demands.* San Francisco: Jossey-Bass.

Burke, Joseph C., and Henrik Minassians. 2003. *Performance reporting: "Real" accountability or accountability lite. Seventh annual survey.* Albany: Rockefeller Institute of Government.

Burke, Joseph C., and Andreea M. Serban, eds. 1998. *Performance funding for public higher education: Fad or trend?* New Directions for Institutional Research #97. San Francisco: Jossey-Bass.

Business–Higher Education Forum. 2004. *Public accountability for student learning in higher education.* Washington, DC: American Council on Education.

Caplow, Theodore, and Reece J. McGee. 1958. *The academic marketplace.* New York: Basic Books.

Carini, Robert M., George D. Kuh, and Stephen P. Klein. 2006. Student engagement and student learning: Testing the linkages. *Research in Higher Education* 47:1–32.

Chatman, Steve. 2007. *Institutional versus academic discipline measures of student experience: A matter of relative validity.* Berkeley: Center for Studies in Higher Education. CSHE 8.07 (May).

Chickering, Arthur W., and Zelda Gamson. 1987. Seven principles for good practice in undergraduate education. *AAHE Bulletin* 39:3–7.

Clark, Burton R. 1987. *The academic profession: National, disciplinary, and institutional settings.* Berkeley and Los Angeles: University of California Press.

Commission on the Future of Higher Education. 2006. *A test of leadership: Charting the future of U.S. higher education.* Washington, DC: Commission on the Future of Higher Education. A Report of the Commission Appointed by Secretary of Education Margaret Spellings.

Cornwell, Grant, and Eve Walsh Stoddard. 1999. *Globalizing knowledge: Connecting international and intercultural studies.* Washington, DC: Association of American Colleges and Universities.

Cremin, Lawrence. 1961. *The transformation of the school, 1876–1957.* New York: Random House.

Cummings, William S., and Martin J. Finkelstein. 2007. The changing academic profession in the United States. Unpublished paper presented at the Hiroshima Conference on the Changing Academic Profession.

DeAngelo, Linda, Sylvia Hurtado, John H. Prior, Kimberly R. Nelly, Jose Luis Santos, and William S. Korn. 2007. *American college teacher: National norms for the 2007–2008 HERI Faculty Survey.* Los Angeles: Higher Education Research Institute.

Dey, Eric L., Claudia E. Ramirez, William S. Korn, and Alexander W. Astin. 1993. *The American college teacher: National norms for the 1992–93 HERI Faculty Survey.* Los Angeles: Higher Education Research Institute, Graduate School of Education, UCLA.

Dougherty, Kevin J., and Rebecca S. Natow. 2009. The political origins of state-level performance funding for higher education. Unpublished paper, Teacher's College, Columbia University.

Dougherty, Kevin J., and Monica Reid. 2007. *Fifty states of achieving the dream: State policies to enhance access to and success in community colleges across the United States.* New York: Teacher's College, Columbia University.

Dwyer, Carol A., Catherine M. Millet, and David G. Payne. 2006. *A culture of evidence: Post-secondary assessment and learning outcomes.* www.ets.org (accessed November 16, 2008).

Education Commission of the States. 1983. *Report of the task force on education and economic growth.* Denver: Education Commission of the States.

Educational Testing Service. 2003. *Quality, affordability and access: Americans speak of higher education.* Princeton: ETS. www.ets.org/media/2003/report.pdf (accessed January 4, 2007).

Ewell, Peter T. 2001a. *Accreditation and student learning outcomes: A proposed point of departure.* Washington, DC: Council on Higher Education Accreditation.

———. 2001b. Statewide testing in higher education. *Change* 33 (March-April): 50–70.

———. 2004. *Accreditation and the provision of additional information to the public about institutional and program performance.* Washington, DC: Council for Higher Education Accreditation.

———. 2005. Can assessment serve accountability: It depends on the question. In *Achieving accountability in higher education: Balancing public, academic, and market demands*, ed. Joseph C. Burke and Associates. San Francisco: Jossey-Bass.

Ewell, Peter T., and Margaret A. Miller. 2005. *Measuring up on college-level learning.* San Jose, CA: National Center for Public Policy in Higher Education.

Freidson, Eliot. 2005. *Professionalism: The third logic.* Chicago: University of Chicago Press.

Galvez, Antonio, Mercedes Maqueda, Manuel Martinez-Bueno, and Eva Valdivia. 2000. Scientific publication trends in the developing world—what can the volume and authorship of scientific articles tell us about scientific progress in various regions? *American Scientist* 88:526–33.

Geiger, Roger L. 2004. *Knowledge and money: Research universities and the paradox of the marketplace.* Stanford: Stanford University Press.

Gilman, David A. 1978. The logic of minimal competency testing. *NASSP Bulletin* 62:56–63.

Glassick, Charles M., Mary Taylor Huber, and Gene I. Maeroff. 1997. *Scholarship assessed: A special report on faculty evaluation.* San Francisco: Jossey-Bass.

Golde, Chris M., and Timothy M. Dore. 2001. *At cross purposes: What the experiences of doctoral students reveal about doctoral education.* www.phd-survey.org (accessed August 4, 2006). Report prepared for the Pew Charitable Trusts.

Grigsby, Mary. 2009. *Life through the eyes of students.* Albany: State University of New York Press.

Hall, Roberta M., and Bernice R. Sandler. 1982. *The classroom climate: A chilly one for women?* Washington, DC: Association of American Colleges.

Hart, Peter D., and Associates. 2006. *How should colleges prepare students in today's global economy?* Washington, DC: Peter D. Hart Research Associates.

Hawthorne, Joan. 2008. Accountability and comparability: What's wrong with the VSA approach? *Liberal Education* 94 (2): 24.

Horta, Hugo, and Francisco Veloso. 2007. Opening the box: Comparing EU and US scientific output by scientific field. *Technological Forecasting and Social Change* 74:1334–56.

Huber, Mary Taylor, and Patricia Hutchings. 2005. *The advancement of learning: Building the teaching commons.* San Francisco: Jossey-Bass.

Hutchings, Patricia, and Lee S. Shulman. 1999. The scholarship of teaching: New elaborations, new developments. *Change* 31 (September/October): 10–15.

Jaschik, Scott. 2009. Out of the loop. www.insidehighered.com/news/2009/06/12/survey (accessed August 31, 2009).

Jencks, Christopher, and David Riesman. 1968. *The academic revolution.* New York: Doubleday Anchor.

Johnson, Valen E. 2003. *Grade inflation: A crisis in college education.* New York: Springer.

Kerr, Clark. 1963. *The uses of the university.* Cambridge, MA: Harvard University Press.

Kuh, George D. 2009. The National Survey of Student Engagement: Conceptual and empirical foundations. In *Using student engagement data in institutional research, new directions for institutional research,* no. 141, ed. Robert Gonyea and George D. Kuh. San Francisco: Jossey-Bass.

Kuh, George D., and Stanley Ikenberry. 2009. *More than you think, less than we need: Learning outcomes assessment in American higher education.* Champaign, IL: National Institute for Learning Outcomes Assessment.

Kuh, George D., Jillian Kinzie, John H. Schuh, and Elizabeth J. Whitt. 2005. *Assessing conditions to enhance educational effectiveness: The inventory for student engagement and successes.* San Francisco: Jossey-Bass.

Ladd, Everett C., Jr. 1979. The work experience of American college professors: Some data and an argument. In *Current issues in higher education.* San Francisco: Jossey-Bass.

Lattuca, Lisa R., Patrick T. Terenzini, and J. Fredericks Volkwein. 2006. *Engineering change: A study of the impact of EC 2000.* Baltimore: ABET, Inc.

Lesgold, Alan, Michael J. Feuer, and Allison M. Black, eds. 1997. *Transitions in work and learning.* Washington, DC: National Academy Press. U.S. Department of Education, Office of Educational Research and Improvement.

Lindholm, Jennifer A., Alexander W. Astin, Linda J. Sax, and William S. Korn. 2002. *The American college teacher: National norms for the 2001–02 HERI Faculty Survey.* Los Angeles: Higher Education Research Institute, Graduate School of Education, UCLA.

Lindholm, Jennifer A., Katalin Szelenyi, Sylvia Hurtado, and William S. Korn. 2005. *The American college teacher: National norms for the 2004–2005 HERI Faculty Survey.* Los Angeles: Higher Education Research Institute, Graduate School of Education, UCLA.

Lusterman, Seymour. 1977. Education and industry. Report No. 719. New York: Conference Board.

McLendon, Michael K., James C. Hearn, and Russell Deaton. 2006. Called to account: Analyzing the origins and spread of state peformance-accountability policies for higher education. *Educational Evaluation and Policy Analysis* 28:1–24.

Meizlish, Deborah, and Matthew Kaplan. 2008. Valuing and evaluating teaching in academic hiring: A multi-disciplinary, cross-institutional study. *Journal of Higher Education* 79:489–512.

Miller, Richard I., Hongyu Chen, Jerome B. Hart, and Clyde B. Killian. 1990. New approaches to faculty evaluation—a survey, initial report. Athens, OH: Ohio University. Report submitted to the Carnegie Foundation for the Advancement of Teaching.

Millett, Catherine M., Leslie M. Stickler, David G. Payne, and Carol A. Dwyer. 2007. *A culture of evidence: Critical features of assessments for postsecondary student learning.* Princeton: Educational Testing Service.

Moffatt, Michael. 1989. *Coming of age in New Jersey: College and American culture.* New Brunswick: Rutgers University Press.

Musil, Carol M., ed. 1992. *The courage to question: Women's studies and student learning.* Washington, DC: Association of American Colleges and Universities.

Nathan, Rebekah. 2005. *My freshman year: What a professor learned by becoming a student.* Ithaca: Cornell University Press.

National Commission on Accountability in Higher Education. 2005. *Accountability for better results: A national imperative for higher education.* Washington, DC: State Higher Education Executive Officers.

National Governors Association. 1986. *A time for results: The governors' 1991 report on education.* Washington, DC: National Governors Association.

National Institute of Education Study Group on the Conditions of Excellence in American Higher Education 1984. *Involvement in learning: Realizing the potential of American higher education.* Washington, DC: National Institute of Education.

National Research Council (NRC). 2005. *Measuring literacy: Performance levels for adults.* Washington, DC: National Academies Press.

National Science Foundation, Division of Science Resources Statistics. 2007. *Changing U.S. output of scientific articles 1988–2003,* Derek Hill, Alan I. Rapaport, Rolf F. Lehming, and Robert K. Bell. Arlington, VA: National Science Foundation.

National Survey of Student Engagement (NSSE). 2000. *The NSSE 2000 Report: National benchmarks of effective educational practice.* Bloomington, IN: Indiana University, Center for Postsecondary Research.

———. 2008. *Promoting engagement for all students: The imperative to look within. 2008 results.* Bloomington, IN: Indiana University, Center for Postsecondary Research.

Osborne, David, and Ted Gaebler. 1992. *Reinventing government: How the entrepreneurial spirit is transforming the public sector.* New York: Penguin Books.

Pascarella, Ernest T., Tricia A. Seifert, and Charles Blaich. 2009. Validation of the NSSE benchmarks and deep approaches to learning against liberal arts outcomes. Paper presented at the annual meeting of the Association for the Study of Higher Education. Jacksonville, FL (November).

Reinvention Center. 2002. *Reinventing undergraduate education: Three years after the Boyer Report.* Stony Brook: Reinvention Center.

Riesman, David. 1980. *On higher education: The academic enterprise in an era of rising student consumerism.* San Francisco: Jossey-Bass.

Sax, Linda J., Alexander W. Astin, Marisol Arrendondo, and William S. Korn. 1996. *The American college teacher: National norms for the 1995–96 HERI Faculty Survey.* Los Angeles: Higher Education Research Institute, Graduate School of Education, UCLA.

Sax, Linda J., Alexander W. Astin, William S. Korn, and Shannon K. Gilmartin. 1999. *The American college teacher: National norms for the 1998–99 HERI Faculty Survey.* Los Angeles: Higher Education Research Institute, Graduate School of Education, UCLA.

Schaar, Jack, and Sheldon Wolin. 1965. A special supplement: Berkeley and the fate of the multiversity. *New York Review of Books* (March 11). www.nybooks.com/articles/13005 (accessed August 30, 2009).

Schuster, Jack H., and Martin Finkelstein. 2008. *The American faculty: The restructuring of academic work and careers.* Baltimore: Johns Hopkins University Press.

Shavelson, Richard J. 2007. *A brief history of student learning assessment: How we got where we are and a proposal for where to go next.* Washington, DC: Association of American Colleges and Universities.

Shulenburger, David. 2008. *Measuring core educational outcomes at research universities for improvement and accountability.* Washington, DC: National Association of State Colleges and Land-Grant Universities. Paper prepared for the October 2008 seminar on Measuring Undergraduate Learning Outcomes: A Working Agenda for Public Research Universities.

Shulman, Lee S. 1989. Toward a pedagogy of substance. *AAHE Bulletin* (June).

———. 1993. Teaching as community property: Putting an end to pedagogical isolation. *Change* (November/December): 6–7.

———. 1997. Professing the liberal arts. In *Education and democracy: Re-imagining liberal learning in America*, ed. Robert Orrill. New York: College Board Publications.

Shulock, Nancy, and Colleen Moore. 2002. *An accountability framework for California higher education: Informing public policy and improving outcomes.* Sacramento: Center for California Studies.

Squires, Gregory D. 1979. *Education and jobs: The imbalancing of the social machinery.* New Brunswick, NJ: Transaction Books.

Tuchman, Gaye. 2009. *Wannabe U: Inside the corporate university.* Chicago: University of Chicago Press.

Zemsky, Robert, and Maria Iannozzi. 1998. *A reality check: First findings from the EQW National Employer Survey.* ERIC #ED382811. EQW Issues No. 10.

Zumeta, William. 2001. Public policy and accountability in higher education: Lessons from the past and present for the new millennium. In *The States and public higher education policy: Affordability, access, and accountability*, ed. Donald E. Heller, 155–97. Baltimore: Johns Hopkins University Press.