History of Education Quarterly (2021), **61**, 546–550
doi:10.1017/heq.2021.42

CAMBRIDGE
UNIVERSITY PRESS

REVIEW ESSAY

Scott M. Gelber. Grading the College: A History of Evaluating Teaching and Learning

Baltimore: Johns Hopkins University Press, 2020. 248 pp.

Jonathan Zimmerman. The Amateur Hour: A History of College Teaching in America

Baltimore: Johns Hopkins University Press, 2020. 320 pp.

Daniel A. Clark

Indiana State University
Dan.Clark@indstate.edu

Although Jonathan Zimmerman's *The Amateur Hour: A History of College Teaching in America* and Scott M. Gelber's *Grading the College: A History of Evaluating Teaching and Learning* cover much of the same historical terrain, the focus of each book proves quite distinct and yet highly complementary. Zimmerman best articulates the stark paradox at the center of both studies. As the United States developed the most elaborate system of higher education and the scholarly endeavors of faculty members evolved into a "highly professionalized enterprise, marked by elaborate codes of credentialing and practice" (p. 10), college teaching in contrast remained a game of chance, conducted by "amateurs, working according to folkloric traditions rather than codified ones" (p. 226). Zimmerman's work attempts a simple and yet daunting task: to supply a history of college teaching in America. Zimmerman excels in discussing the stories of great lecturers and efforts for reform, but often leaves more questions than answers when it comes to the myriad attempts (usually dismissed by faculty) to systematically assess college teaching, exactly where Gelber's work fills in the gaps. As Gelber notes in his conclusion, the evaluation of collegiate teaching and learning "ranks among the greatest unsolved problems of academia—a basic scholarly dilemma" (p. 156). Taken together, these two books may help academicians begin to work toward solutions, simply by finally offering accessible historical overviews of their related areas, though Gelber seems the more optimistic of the two with regard to the possibility (or necessity) of positive change.

Zimmerman's study is the more straightforward in laying out a historical narrative, and his first two chapters deftly explore the key transformation of American higher education (the rise of the American university and its scholarly ideals) that fundamentally changed the college experience and its pedagogical traits. He first examines the nineteenth-century small college, where he finds no shortage of accounts of

© History of Education Society 2021

https://doi.org/10.1017/heq.2021.42 Published online by Cambridge University Press

exemplary professors. Yet Zimmerman also delights in relaying the shortcomings of the recitation, the central pedagogical method of that era. In the right hands, the recitation could yield a wonderful Socratic exchange. Too often, however, faculty would merely call on students to recite passages from memory—hardly halcyon days. Then a trickle of scholars trained in German universities began introducing the lecture to emulate their German professors and showcase their new learnedness, part of the well-known story of the professionalization of the American academy, a story Zimmerman admits is hardly novel but which has never been told from the perspective of its impact on collegiate instruction, and his is the best account that I have read on the subtleties related to the recitation and its replacement by the new lecture method. While the hated quizzing via recitation slowly died, the lecture brought on a new problem—boring and listless delivery and mere passive listening. To be sure, gifted lecturers inspired students. But Zimmerman details how the growing faculty devotion to scholarship scuttled any incentive to improve teaching. The coin of the realm for judging faculty from then on, despite many half-hearted attempts at reform, would remain scholarship. Universities and those wishing to emulate them expended no energy to prepare faculty to teach, a "curious blend of system [scholarship] and anarchy [teaching], of expertise and amateurism" (p. 33).

The next six chapters make up the bulk of Zimmerman's study. Three blocks, each consisting of two chapters (one offering an overview, the other detailing reform efforts) explore the interwar years, the Cold War era, and finally the upheaval of the 1960s and 1970s. Zimmerman describes the familiar tandem stories of the growth of mass higher education and the embrace of the scholarly research ideal, but with a focus on its impact on college teaching. In each era, the prevalence of the lecture increased owing to its efficiency and cost-effectiveness as enrollments grew. While the faculty's attention to scholarly output often exacerbated the problem of uninspiring lectures, many faculty and administrators wrestled with the issue of poor teaching as students became ever more assertive in their complaints. Institutions largely proved unwilling, however, to confront the central cause (devotion to scholarship in training and promotion). Having visited an impressive fifty-nine college and university archives, Zimmerman himself notes the repetitive nature of most reform ventures—committees convened, statements were issued, and most professors ignored them and proceeded to teach in the same fashion. Nevertheless, his archival work does yield excellent accounts of often neglected reform efforts, which form the key contribution of his work. He provides admirably detailed descriptions of reforms from the well-known, such as the Harvard tutorial or experimental colleges like Rollins (with its conference system), to the more obscure endeavors, such as B. F. Skinner's teaching machines and the T-group craze of the 1960s and 1970s (emulating encounter groups), all of which tried to counter the mass lecture's apparent deficiencies with more intimate encounters and/or student autonomy, with almost all of them failing. Foundation grants funded nearly all such reform ventures at one point or another, including serious efforts to alter the other end of the problem—faculty training, through promoting pedagogical seminars in graduate programs as well as ultimately championing the teaching-focused Doctor of Arts degree to supplant the research-centered PhD, none of which produced lasting change.

https://doi.org/10.1017/heq.2021.42 Published online by Cambridge University Press

A long epilogue explores higher education teaching reform since 1980, a period when the lecture method came under its most recent assault via pushes for student-centered and active learning concepts, much of which Zimmerman recognizes as old wine in new bottles. Here Zimmerman finally reveals his own perspective that "teaching is a deeply personal and even spiritual act that defies rational organizing" (p. 234), a view perhaps reinforced by his meticulous research, as chapter after chapter quotes faculty, deans, and presidents making the very same arguments justifying why movements to define and measure (and alter) college teaching consistently failed to gain traction.

At several points in Zimmerman's work the evaluation of teaching or judging the effectiveness of a teacher arises, but he dismisses those instances largely because his faculty and administrative sources shared the same vision of teaching as Zimmerman —teaching is a natural talent or aspect of personality that defies rational metrics. One consistently feels that such dismissals mask a substantive tangential history, and that is where Gelber picks up the story. Gelber might not have combed through as many campus archives as Zimmerman, but his mastery of the source material remains impressive. He clearly visited key institutions that pioneered evaluation techniques (the universities of Chicago, Minnesota, and Washington, for instance) and the organizations that supported and led (and at times opposed) such efforts (the Carnegie Foundation for the Advancement of Teaching, the American Association of University Professors, the American Council on Education, and the College Entrance Examination Board). Much of Gelber's sources, though, come from published reports and articles, and he seems to have read practically everyone for the last eighty years, as evidenced by his packed endnotes.

Gelber states plainly that one of the principal outcomes of his research involves debunking the modern faculty perspective that opposes seemingly all attempts at the evaluation of teaching, with faculty judging them all as external political assaults on higher education. His final section on "accountability" covers the genesis of this jaundiced faculty view (one that he sees as partially justified), which emerged after the backlash against higher education in the 1970s and has gained steam since the 1980s with the advent of the "culture wars" and the worship of market-oriented metrics fueling calls for applying outcomes-based assessments for colleges and universities. The bulk of his work, however, "complicates the notion that evaluation was imposed upon passive faculty from the outside" (p. 156), with his research revealing that ever since the movement for educational measurement first emerged in the 1920s, a significant minority of faculty played key roles in forwarding every facet of evaluation endeavors. The two chapters constituting the first section examine attempts to evaluate the teaching of faculty in order to get beyond casual metrics in the form of hearsay and gossip to more rigorous concepts such as peer review and finally student evaluations, the latter being accepted largely in the wake of student protests of the 1960s owing to their ease of administration and the research that found student evaluations rating more consistently effective than even peer review. The second section on the assessment of student learning offers the best overview on the topic that I have seen. Chapter 3 explores the quest to move beyond grades and to use tests to truly establish sound measurement, discussing the formation of the Educational Testing Service from the merger of the American Council on

https://doi.org/10.1017/heq.2021.42 Published online by Cambridge University Press

Education, the College Entrance Examination Board, and other entities, which pioneered testing efforts. Chapters 3 and 4 also offer an excellent account of the development of student surveys to assess hard-to-quantify but prized aspects of the collegiate experience, such as a broadened perspective, with heavy faculty input and cooperation in structuring such surveys. Similarly, his look at the history of accrediting bodies (chapter 5) highlights that such agencies had long been the creatures of the institutions they regulated, and thus were imbued with an academic culture that overall allowed institutions to self-assess and that particularly loathed the idea of statistics comparing institutional student outcomes.

No reader can walk away from Gelber's study without a curious mix of respect and exasperation. Respect, for his research divulging the long and persistent efforts (covering essentially one hundred years) to measure college teaching and curricular effectiveness. Exasperation, for his sobering overview and the light he sheds on the disturbing lack of any distinct progress in addressing basic questions about teaching and learning assessment, despite decades of attention. I appreciated Gelber's dive into the details of such efforts as the University of Minnesota's attempts to use standardized testing to assess their ground-breaking general education curriculum in the 1930s and 1940s, an era when standardized tests to measure college and departmental effectiveness abounded. (Comprehensive department exams were all the rage in that period.) Like many such endeavors, they yielded scant evidence of superior or inferior teaching methods and were abandoned as too time-consuming, disliked by students and faculty. Similarly, his exploration of the history behind the embrace by many faculty of portfolios of student work, judged through rubrics, also proved illuminating. Researchers and faculty counterparts have expended a great deal of time and energy trying to evaluate teaching and learning, and yet even today little consensus has emerged regarding what constitutes good teaching, or how we can quantitatively answer seemingly basic questions such as whether college teaching has improved or declined over time. Such uncertainty only heightens people's exasperation.

Gelber's last chapter examines elements of this frustration as he charts the last forty years in the quest to evaluate collegiate teaching and learning. Various publics (parents and state and federal governments) placed increasing pressure on American higher education to demonstrate its effectiveness, a natural reaction as Americans spent ever more money on tuition. Calls for tracking and ranking institutions on their "outcomes," such as graduation rates and alumni earnings, have grown common. Gelber does a nice job of simultaneously following the growth of faculty defensiveness—how the academy came to forget its deep involvement with evaluation efforts and to bridle at calls for accountability. In this final chapter Gelber also performs the vital function of surveying the last forty years of research on teaching and learning evaluation that overall corroborates much of the faculty's resentment, highlighting again the problems with student evaluations (the privileging of White males over females and minorities, for example) and the obvious shortcomings involved with comparing institutional outcomes without accounting for student origins (in terms of class, race, educational preparation). When trying to assess student outcomes, researchers have not been able to separate out the demographic benefits and liabilities of students from evidence of curricular or institutional effectiveness.

https://doi.org/10.1017/heq.2021.42 Published online by Cambridge University Press

Each author marshals prodigious research toward elucidating a history familiar to most in higher education yet still elusive and controversial. Gelber derives more lessons from his study than Zimmerman, and presents them candidly in a brief conclusion. He advises the decoupling of accountability measures (a necessary evil) from ongoing efforts to evaluate teaching and learning with the aim of improving college teaching, a more cautiously optimistic stance than Zimmerman's resignation regarding the seemingly inscrutable mysteries of college instruction. Each may stand on their own, though I would recommend reading them in tandem. Zimmerman moves in a linear, chronological fashion, while Gelber jumps back and forth in time, even within the same chapter, to follow the rise and fall of various evaluation ventures. It helped in reading Gelber to have Zimmerman's broader treatment as context. Scholars of American higher education will appreciate both these fine works. Just do not expect a happy ending.

https://doi.org/10.1017/heq.2021.42 Published online by Cambridge University Press