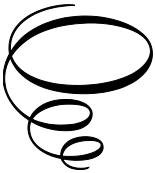
Conflicting Stakeholders in Higher Education Assessment

Conflicting Stakeholders in Higher Education Assessment

By

H. Russell Searight

Cambridge Scholars Publishing



Conflicting Stakeholders in Higher Education Assessment

By H. Russell Searight

This book first published 2024

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN (10): 1-0364-0658-X
ISBN (13): 978-1-0364-0658-5

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

The assessment movement in higher education began in the mid-1980s. Influenced by the broader accountability movement within the professions, the development of educational standards in K-12 education, and the increasing federal funding devoted to higher education, this movement has significantly impacted universities. Today, almost all universities in the United States and many in Europe have established protocols and templates for assessing programs, concentrations or majors, and specific courses. In Great Britain, this movement is known as "Quality Assurance." In both the United States and Europe, accrediting bodies oversee universities and colleges to ensure they meet quality standards. In many European countries, these accrediting bodies are closely associated with government offices of education. In the United States, accrediting bodies have traditionally functioned relatively independently, but recent years have seen increased government scrutiny.

Assessment in higher education involves multiple definitions but commonly consists of establishing educational objectives for students, classes, programs (majors), and the university. The assessment process has recently been extended to co-curricular campus activities, including student services. Evaluations of the extent to which these objectives are met is periodically conducted, and data are gathered, analyzed, and reported (Banta & Palomba, 2015). In theory, these data are fed back into the university system to improve the students' educational experience. Assessment professionals speak the language of outcomes and objectives, which are typically quantitatively measured and reported.

Assessment Stakeholders

Higher education assessment is a form of program evaluation involving multiple stakeholders, including individual faculty, university administrators, professionals responsible for overseeing accreditation and assessment, accrediting bodies, and state and federal government agencies. Although

students historically were intended to receive the benefits of accreditation and assessment, their involvement has been relatively minimal in the US, although this is beginning to change.

As expected, there is considerable disagreement among stakeholders regarding the value of assessing higher education, the goals or objectives of higher education itself, and the methods for conducting assessments. Each stakeholder will likely have a unique perspective on the relevant dimensions for evaluation, how these dimensions should be measured, and the reporting procedures that should be employed.

Accreditation typically focuses on issues such as academic quality, faculty credentials, offices protecting civil rights, and attention to cultural diversity. More recently, likely a response to government scrutiny of private institutions that suddenly close and file for bankruptcy, accreditors have been giving increased attention to financial stability. However, historically, accrediting bodies have not emphasized graduation rates, the cost of obtaining a degree from a specific institution, or post-bachelors admission to graduate and professional schools. Nevertheless, as will be discussed in upcoming chapters, accrediting bodies are now paying greater attention to the financial status of universities and outcomes such as graduation rates.

At the faculty level, there has been a significant emphasis on developing course learning outcomes, which are specific objectives defined in measurable terms for each course. At the end of every semester, data on whether students have achieved these objectives are reported in a central institutional database using proprietary templates that provide an efficient overview of teaching and learning effectiveness. These data are important for accrediting bodies during their periodic reviews of universities. However, an exclusive focus on course learning outcomes may not fully capture the holistic picture of student success and the broader goals of higher education. It is essential to consider other factors, such as developing critical thinking, problem-solving, and communication skills, as well as students' overall well-being and personal growth. Only by embracing a more comprehensive approach to assessment can we ensure that higher education fulfills its fundamental mission of preparing students for a successful future.

As the reader progresses through the book, they may be surprised by the introduction of concepts from philosophy, industrial psychology, sociology, and political science into the discussion of assessment. However, these fields are crucial for understanding how assessment has evolved and the

underlying implications of course and program assessment. As a university instructor, I firmly believe it is essential to move beyond merely describing phenomena to explaining their causes. As a psychologist, I understand the need to examine the impact of the global interconnected community and go beyond the individual to consider broader societal factors.

In the case of higher education assessment, this context includes the history of higher education, accreditation processes, trends in measuring educational outcomes, and political, sociological, and economic factors. For example, assessing students and their learning outcomes in university courses without considering the 12 to 13 years spent in America's primary and secondary education system is an artificial, historical exercise. As I discuss later, universities are not isolated, hermetically sealed, self-contained worlds that exist apart from larger society. To truly understand the purpose and impact of higher education, we must consider the broader social, economic, and political context in which it exists.

The Social Context of Assessment

Structural Inequality

As the number of high school graduates who pursue college and university education continues to increase, so does the representation of students from minority backgrounds and lower-income households who have historically experienced lifelong structural inequalities. Structural inequalities refer to a system where some groups in society enjoy privilege and advantage, which is sustained by broader institutional structures such as the government, legal system, and economic disparities. When policies and practices prevent some members of society from accessing quality education, employment, housing, and other resources, they perpetuate structural inequality (Amadeo, 2022).

The rise in income inequality has exacerbated other structural factors that limit access to resources, resulting in a diminished quality of life. For instance, among primary and secondary education students, those from lower-income neighborhoods had lower math scores on average. This disparity can be attributed to the neighborhood's economic status which accounts for almost 40% of the differential in Math scores (Michigan State University, 2022).

Among minority groups, structural violence takes the form of being subjected to daily microaggressions. These sometimes subtle forms of communication are directed toward students from minority groups and have

been well-documented throughout American educational institutions, including colleges and universities (Ogunyemi et al., 2020).

Furthermore, establishing measured success through students' achievement of course objectives, applying rubrics, standardized tests, and other forms of assessment occurs in a social vacuum that presents itself as value-neutral. However, this approach denies the reality of educational inequality in the United States and other Western countries.

A Historical Example of Unsuccessful Academic Remediation

In the mid-1960s, several government programs were established to address economic, educational, and employment inequality as part of President Johnson's Great Society initiatives. One such program was Head Start, initially directed as a preschool program for lower-income children. It was based upon the finding that lower-income children began elementary school with academic disadvantages relative to their middle-class peers. Head Start was intended to provide preschool children with intensive pre-elementary education, expecting them to enter school at a comparable academic level as children from more advantaged backgrounds (Zigler & Styfco, 2010). Head Start continues and is now a permanent fixture of preschool education in the United States.

However, while Head Start was associated with several positive outcomes, academic success in elementary school was not one of them. Despite the gains demonstrated by Head Start graduates in reading and language skills, students who had been in the program did not enter elementary school at the same academic level as students from more economically advantaged backgrounds. Furthermore, after approximately two years into elementary school, the gains demonstrated by Head Start graduates disappeared (Smith & Bissell, 1970; Westinghouse Learning Corporation, 1969). While there may be multiple reasons for this sobering outcome, evaluators suggested that children receiving Head Start were more likely to continue their education in poor-quality elementary schools, an obstacle that the benefits of Head Start could not overcome.

Higher Education and Elitism

Finally, with the surge in high school graduates continuing their education, some voices have criticized the opening of higher education to a much broader audience, claiming that it diminishes its quality (Mintz, 2022). While there are valid concerns about the academic background provided by

our primary and secondary school systems, some of the labels placed on underprepared minority and first-generation students have an elitist tone. For example, some discussions of first-generation college students' deficient social capital border on condescension (Searight, 2024).

In recent years, I have been involved in several projects focusing on first-generation university students. There are various definitions of first-generation students, but one of the more common ones is an incoming university student from a family in which neither parent has a bachelor's degree. Universities have developed special offices and programs devoted to first-generation students. These "interventions" often address academic weaknesses and less tangible deficits, including social capital. However, placing these students in a "special" category that focuses on their attributed deficits can be potentially harmful (Cornell & Searight, 2023).

It is important to acknowledge the challenges that first-generation students may face in navigating higher education, but it is equally important to recognize their strengths and potential. Instead of treating them as a homogenous group with deficits, universities should provide support and resources that acknowledge their diverse experiences and backgrounds. By doing so, universities can empower these students and help them realize their full potential. (Ives et al., 2020). However, this perspective should not delegitimize the many strengths that first-generation students possess, including a strong sense of pride, excellent time management skills, and the ability to balance multiple responsibilities, such as work, family obligations, and a full course load. These students' persistence and ability to overcome obstacles are also noteworthy, as is their loyalty to their families and communities (Cornell & Searight, 2023; Searight, 2024). Unfortunately, these character strengths are often not recognized in the deficit-oriented descriptions of first-generation students.

It is essential to recognize that the White Northern European model of child-rearing, which values independence, achievement, and self-direction, may not be congruent with the values of many minority and first-generation students. These students often prioritize a sense of duty and obligation to important others while pursuing higher education. The value of social solidarity and collectivism that guides many of these students may conflict with the predominantly White European value of individual achievement that predominates among university administrators (Stephens et al., 2012; Cornell & Searight, 2023). It is essential to acknowledge and celebrate the diversity of values, experiences, and strengths that first-generation students bring.

Conclusion

While my discussion of higher education assessment may come across as critical or biased, I am not opposed to assessment itself. I believe that continuous assessment of students and myself as an instructor can improve the quality of higher education. Meaningful assessment can help me in fulfilling my responsibility to provide high-quality instruction. Teaching is a calling for me, and I take it seriously.

While I do not suggest abandoning assessment entirely, I believe that some of the time, money, and energy currently devoted to assessment could be redirected to faculty development to improve teaching (Eubanks et al., 2020). In conjunction with student input, reflective self-criticism is essential to quality teaching. In conclusion, I am open to assessment and its potential benefits, but I also believe that we should strive to find a balance that prioritizes assessment and faculty development to ensure the best possible education for our students.

The 35–40-year history of assessment in higher education has not led to improved student educational outcomes. As a bureaucratized exercise in quasi-scientific measurement, higher education assessment, while providing reassurance to university accreditors that we are productively examining the effectiveness of classes, programs (majors), and, more recently, co-curricular programming, has not fulfilled the stated mission of improving student learning. Additionally, despite the time and money devoted to assessment, the practice has not demonstrably led to improved pedagogy and for many faculty members, is seen as a poor use of professional time (Eubanks, et al. 2021; Gilbert, 2018)

In later chapters, I will discuss the current state of incoming university students' academic skills and the outcomes upon receiving a bachelor's degree. With the rapid increase in the college student population in the United States and Western Europe and the corresponding diversity of these students, some important social and economic trends undoubtedly impact educational outcomes.

At the end of every semester, I receive feedback from my students regarding my performance as an instructor. This feedback is presented in two forms: numerical ratings and narrative comments. The numerical ratings range from 1 through 5 and cover multiple dimensions such as clarity of communication, approachability, and the promptness of returning graded tests. However, I have found these end-of-semester ratings to be of limited

help in improving my teaching. Although the ratings are usually summarized as mean ratings and ranges and entered into a spreadsheet, it is difficult to interpret these statistics. For example, a rating of 3.5/5 on the dimension "The instructor was enthusiastic" does not provide specific direction for improvement.

On the other hand, the narrative comments I receive are constructive. Students respond to pre-established prompts, such as "What did your instructor do that you really appreciate?" or "What advice or information do you want to share with your instructor to help them improve their course?" These comments provide specific directions on what I could change as an instructor, including my interactions with students, testing methods, textbook choices, and pedagogical strategies.

Overall, while numerical ratings provide a general sense of how students perceive my teaching, narrative comments are more helpful in providing specific feedback on how I can improve as an instructor. Therefore, I prioritize the narrative comments and use them as a basis for making changes to my teaching methods in the future.

Sometimes, as faculty members, we may feel irritable (i.e., "This is a major waste of my time.") as we enter our assessment data, and may be tempted to blame deans, assessment coordinators, members of accrediting bodies, and other administrators for requiring us to complete these actuarial exercises. However, it is essential to understand the broader context in which these requirements occur. Assessment coordinators are not trying to make faculty members' lives miserable. Instead, they are responding to the same social and historical climate that has led to the current preoccupation with accountability in higher education and all professions, including health care.

In many respects, all of us, including our students, have had to respond to the demands of measurement-focused accountability. Our students are quite familiar with these practices since before they even entered college, they were subjected to standardized tests on an annual basis beginning in elementary school. These high-stakes tests, even given to second graders, often determine the financial support that their school districts receive.

Therefore, it is essential to understand that the emphasis on assessment and accountability in higher education is not unique or malicious but rather part of a more significant societal trend. Faculty members need to engage with assessment constructively and use it as a tool for improving teaching and

learning. Instead of blaming others, we should work with administrators and assessment coordinators to develop effective assessment practices that benefit both faculty and students. By doing so, we can ensure that our students receive the best possible education and that our institutions remain accountable to the broader society.

I reiterate my support for using assessment meaningfully to improve the quality of teaching, student learning, and program development. As an instructor, I believe engaging in reflective activities is essential to improving my teaching. If I do offer criticism, I feel responsible for presenting viable alternatives.

In the final chapter of this book, I suggest assessment strategies that I hope can meet the requirements of accreditors and state government offices while providing better-quality information. Some of these approaches may seem incomplete, insufficiently rigorous, or impractical. However, it is important to note that members of the assessment profession have acknowledged these limitations of the current system.

This book presents a brief history and background of assessment and accreditation before exploring the purpose of higher education. While this sequence may seem illogical, it reflects an important reality. In the rush to establish measurable outcomes (Layzell, 1998), fundamental questions about higher education's purpose(s) were not given the attention they deserved.

Before proceeding, the author would like to thank Dr. Lynn Gillette, Interim President, and Dr. Kathy Berchem, Dean of the College of Health and Emergency Responders at Lake Superior State University, for supporting the sabbatical year during which the initial draft of this book was written. Additionally, the author deeply appreciates Ms. Madison Anderson's thorough editing and proofreading of this work.

References

Amadeo (2022). What is structural inequality? How structural inequality stifles the American dream. *The Balance.* https://www.thebalancemoney.com/structural-inequality-facts-types-causes-solution-4174727

Banta, T. W., & Palomba, C. A. (2015). *Assessment essentials: Planning, implementing, and improving assessment in higher education*. John Wiley & Sons.

Bissel, J.S. (1970). Report analysis: The impact of Head Start. *Harvard Educational Review, 40*, 51-104.

Cornell, A., & Searight, H. R. (2023). The challenges behind living a double life among first-generation university students: The clash between interdependent and independent identities. In P. Aloka & R. Akuna (Eds.), *Handbook of Research on Coping Mechanisms for First-Year Students Transitioning to Higher Education* (pp. 142-159). IGI Global.

Eubanks, D., Fulcher, K., & Good, M. (2021). The next ten years: The future of assessment practice. *Research & Practice in Assessment, 16*(1).

Gilbert, E. (2018). An insider's take on assessment: It may be worse than you thought. *The Chronicle of Higher Education*, 12.

Ives, J., & Castillo-Montoya, M. (2020). First-generation college students as academic learners: A systematic review. *Review of Educational Research, 90*(2), 139-178.

Layzell, D. T. (1998). Linking performance to funding outcomes for public institutions of higher education: The U.S. experience. *European Journal of Education, 33*(1), 103-111.

Michigan State University. (2022). Education Policy Initiative: K-8 Student Achievement and Achievement Gaps on Michigan's 2020-21 Benchmark and Summative Assessments.

Mintz, S. (2022, February 8). Is college for everyone? *Inside Higher Ed*. Retrieved from [link]

Ogunyemi, A. O., Babalola, S. O., & Olaleye, B. (2020). Psychological determinants of xenophobic behaviours. *Islamic University Multidisciplinary Journal, 7*(2), 208-217.

Stephens, N. M., Fryberg, S. A., Markus, H. R., Johnson, C. S., & Covarrubias, R. (2012). Unseen disadvantage: How American universities' focus on independence undermines the academic performance of first-generation college students. *Journal of Personality and Social Psychology, 102*(6), 1178.

Searight, H.R. (2024, June,14). First generation students in higher education: A cultural mismatch. Paper presented at the 2024 Upper Peninsula Teaching and Learning Conference, Bay College, Michigan.

Westinghouse Learning Corporation. (1969). The impact of Head Start: An evaluation of the effects of Head Start on children's cognitive and affective development. A report presented to the Office of Economic Opportunity, Ohio University.

Worthen, M. (2018, February 23). The misguided drive to measure "learning outcomes." *New York Times*. Retrieved from https://www.nytimes.com/2018/02/23/opinion/sunday/colleges-measure-learning-outcomes.html

Zigler, E., & Styfco, S. J. (2010). *The hidden history of Head Start*. Oxford University Press.

CHAPTER 2

ASSESSMENT, ACCREDITATION AND GOVERNMENT'S ROLE IN HIGHER EDUCATION: A BRIEF HISTORY

Introduction

This chapter offers a brief historical overview of assessment development in higher education and the corresponding evolution of institutional accreditation. Although assessment has become an indispensable feature of colleges and universities, it is doubtful that it would have gained such prominence without the influence of accreditation requirements. While I strive to discuss assessment and accreditation separately, it will soon become apparent that these two concepts are inextricably linked.

Some experts propose a triad that oversees universities, where states primarily focus on consumer protection, the federal government concentrates on financial aid, and accreditors emphasize educational quality. However, this chapter will demonstrate that such categorization significantly oversimplifies the situation. Accreditors have been urged to consider factors like graduation rates and employability, while state governments have increasingly engaged in determining universities' instructional content, particularly around topics like diversity. Moreover, the federal government has periodically criticized accrediting bodies.

Contrary to the assumption that accreditation and college curricula are non-controversial and mundane issues, regional and national news reports suggest a more complex reality (Miller et al., 2023).

Assessment and Accreditation

Program Evaluation and Education

Some of the early leaders in higher education assessment came from the program evaluation movement (Banta & Palmoba, 2015). While not typically recognized as part of academic assessment, educational evaluation began in the 1960s. One of the preeminent voices in program evaluation during the latter part of the 20th century, Michael Scriven, famously said, "Evaluation is a very young discipline-although. It is a very old practice" (Scriven, 1996; p. 395 cited in Hogan, 2007).

In the late 18th century, Wiliam Farish is credited with assessing students' performance and using the resulting data to rank their averages (Hogan, 2007). In 1868, the Powlis Commission published the results of an investigation of student performance in Ireland. In a variation of the relationship between government funding and a school's performance, the Commission recommended annual examinations of students in reading, spelling, and arithmetic to determine teacher salaries (Hogan, 2007). In the mid-1800s, as Secretary of the Massachusetts Board of Education, Horace Mann encouraged mass testing to determine if a student should be promoted to the next grade (Hogan, 2007).

In the mid-1960s, two distinct pathways for program evaluation emerged. The first pathway focused on impact assessment, which became particularly important in the context of Lyndon Johnson's Great Society programs (Madaus et al., 1983). With the implementation of these programs, it became necessary to evaluate their effectiveness and determine if they were achieving their intended goals, which would inform decisions on whether they should receive ongoing funding from Congress.

One area where impact assessment was particularly important was in early childhood educational programs such as Head Start. It was necessary to evaluate these programs to determine if they positively impacted the lives of children and families, particularly those in poverty. To do this, evaluators often relied on quasi-experimental designs, which allowed them to compare the outcomes of those who participated in the program to a comparison group that did not.

The impact approach to program evaluation aimed to measure the effects of a program or intervention on specific outcomes of interest, such as academic achievement or health status. It often involved collecting data on program

participants before and after the intervention and comparing it to a control group. While impact evaluation was not a new concept, it gained prominence during the 1960s as policymakers sought evidence-based approaches to decision-making (Dietz, 1987).

The impact approach in program evaluation often relies on using quasi-experimental designs to address the practical problem of conducting research in real-world settings. Although comparison groups were often used in social interventions, it was necessary to rule out the effects of possible confounding factors such as historical changes, sensitization through repeated testing, and unintended bias in selecting participants (Cook & Campbell, 1986). By systematically listing and ruling out these threats, program evaluators could be confident that program outcomes were due to the intervention itself rather than extraneous factors.

Furthermore, program evaluators could systematically study issues with political significance, such as whether a guaranteed income would reduce employment. By the mid-1970s, program evaluation had become an established profession with its own professional societies and journals (Hogan, 2007). Program evaluation continued to spread into other areas, such as organizational assessment, to assist management, particularly in human service agencies. Additionally, the growing emphasis on training in organizations, such as in-house workshops devoted to diversity, often included evaluating effectiveness (Tennant et al., 2002).

A second program evaluation pathway is termed "performance auditing." Performance auditing is less rooted in the experimental tradition and more significantly influenced by organizational management theory (Hogan, 2007). Auditors tend to examine processes to determine if procedures are being executed as planned and assess whether legal and ethical standards are being met. In some ways, performance auditing shares elements with the management approach known as Total Quality Management (TQM), which focuses on processes rather than outcomes (Dahlgaard et al., 1995) and will be discussed later in this book.

Davis (1990) observes that auditors are more likely to detect fiscal fraud than those conducting summative impact assessments. Performance auditing has been applied to a much broader range of issues beyond human service and educational program evaluation, such as the management of retirement and pension plans (Hogan, 2007).

Assessment in Higher Education

Higher education assessment has primarily developed in response to the demands of accrediting bodies. Additionally, the growing emphasis on consumer satisfaction and the federal government's involvement in higher education through financial aid programs have contributed significantly to the assessment expansion. In the 1980s, the Reagan Administration conducted a study, later titled *A Nation at Risk* (National Commission on Excellence in Education, 1983), highlighting serious concerns about education quality at all levels, including colleges and universities. The report revealed that a considerable percentage of K-12 students' performance was significantly below normative expectations. The document characterized the state of U.S. education with the memorable phrase, "...a rising tide of mediocrity." Concerns about university student learning have persisted and resurfaced more recently in Arum and Roska's 2011 book, *Academically Adrift: Limited Learning on College Campuses.*

The year 1984 marked the beginning of modern higher education assessment. The Southern Association of Colleges and Schools, now known as the Southern Association of Colleges and Schools Commission on Colleges, requested that their affiliated institutions demonstrate effectiveness in addressing student learning. This led to the first national conference on higher education assessment, which took place in South Carolina in 1985 (Ewell, 2002).

At the same time, a competency-based approach was gaining popularity in assessment outside of higher education. This approach focused on determining whether elementary and secondary school students had achieved grade-appropriate knowledge and skills. In business settings and personnel management, the assessment center approach, in which candidates for positions are observed and rated while they carry out simulated work activities, has become increasingly popular (Thornton & Byham, 2013). Today, many assessment centers are available online, evaluating candidates for management positions by having them respond to vignettes that often include workplace problems (Howland et al., 2015). While not a direct application of the assessment center, Alverno College's competence orientation to assessment in the early 1970s is consistent with this approach (Allen, 2016). In addition to eight core competencies: Communication, Analysis, Valuing, Problem Solving, Social Interaction, Developing a Global Perspective, Effective Citizenship, and Global Engagement—all Alverno students were required to complete an off-campus internship. The competencies were developed through the college's

assessment center with input from businesses and professionals in the surrounding community (Allen, 2016).

As a starting point, higher education assessment requires precise definitions. In the mid-1980s, there were three distinct definitions of assessment, each reflecting a different emphasis (Ewell & Cumming, 2017). One definition focused on mastery, representing a competency-based approach that concentrated on directly observing individual performance. The second definition emerged from the accountability movement and highlighted mass testing in the K-12 system (Nodine, 2016), with the analysis centering on students' aggregate performance within specific schools or districts (Nodine, 2016). Lastly, the influence of program evaluation was evident in the third definition, which emphasized collecting data to improve teaching and curricula (Ewell & Cumming, 2017). This final definition most closely approximates current assessment practices in higher education.

Ewell and Cumming (2017) highlight a critical challenge faced during the early days of assessment: the need to rapidly identify credible and effective methods for gathering evidence of student learning. At that time, most existing instruments had been designed for other purposes (Ewell & Cumming, 2017, p.2).

In response to the growing demand for assessment tools and the pressure to implement an active assessment plan, universities initially turned to proprietary testing organizations. These companies developed standardized national examinations such as the ACT Collegiate Assessment of Academic Proficiency and the ETS Major Field Tests (Ewell & Cumming, 2017).

Over the past few decades, the expectations for higher education assessment have become increasingly rigorous (Ewell & Cumming, 2017). By the 1990s, accreditors began to require institutions to have documented assessment plans. By the 2000s, institutions were mandated to provide data on learning and institutional outcomes. Between 2010 and 2020, assessment standards became more complex, requiring demonstration of pedagogical and institutional improvement. Furthermore, the assessment domain expanded, influencing various university activities such as strategic planning, budgeting, and pedagogical design and programs.

By the 1990s, it became increasingly clear that assessment in higher education was not just a passing trend from the 1980s but rather a permanent fixture (Ewell & Cumming, 2017). Establishing assessment as a distinct profession in higher education was one indicator of its enduring nature. The

role of director of assessment, vice president, or dean for assessment became a regular part of university administration. Additionally, some universities began offering graduate degrees with a concentration in educational assessment. In 1993, it was reported that 98% of the institutions participating in the American Council of Education had implemented some form of assessment program, further indicating the widespread adoption and acceptance of assessment practices in higher education.

Higher education assessment is not limited to the United States; it is also prevalent in Europe, where the evaluation of learning outcomes has become increasingly crucial. In Europe, this process is typically called "Quality Assurance," which emphasizes the importance of establishing standards that enable the transfer of courses and recognition by universities throughout the European Union.

Similar to the United States, the number of secondary school graduates entering higher education in European countries has significantly increased over the years, from 5% in the early 20th century to 50-60% currently (OECD, 2024; Osborne, 2003).

One notable difference between the European assessment model and the U.S. is that students have a more influential voice in European higher education. For instance, the European Student Union addresses issues related to quality assessment, ensuring that students' opinions are considered and incorporated into decision-making processes.

Accreditation in Higher Education

The establishment of standards for higher education institutions can be traced back to the mid-1800s when unregulated for-profit commercial schools were on the rise in the United States. Two individuals, Henry Bryant and Henry Stratton, pioneered this trend and developed a large corporate entity with multiple for-profit institutions (Angulo, 2017). These educational conglomerates charged schools and colleges a fee of up to 50% of the tuition collected to become part of their network. Blatant advertising campaigns promised a life of wealth to those who completed short courses in bookkeeping or telegraphy (Angulo, 2017). Degrees were often awarded based on only months or weeks of instruction.

Advocates of these for-profit schools, particularly in business and law, criticized traditional not-for-profit colleges for their emphasis on languages, history, and philosophy, which were deemed irrelevant to the skills required

in industry. By 1895, close to 100,000 students in the U.S. were attending for-profit commercial schools (Angulo, 2017). This increase led to a growing concern about the quality of education and the need to establish standards to ensure that students received a proper education.

Law was the first profession to challenge the for-profit institutions of the time. In 1900, the American Bar Association issued an official statement stating that law schools would not be operated as commercial enterprises. The research university, imported from Germany, became the model for respectable law schools to operate (Angulo, 2017).

Medical education underwent a radical transformation within a relatively short period. In the late 1800s, there was a proliferation of proprietary medical schools competing for students (Miller & Weiss, 2008). These early for-profit institutions led to a relaxation of admission and graduation requirements, which raised concerns about the quality of medical education. In response, the Carnegie Foundation commissioned Flexner's report, which exposed the inadequacies of these schools (Angulo, 2017).

Ten years after Flexner's report was published in 1910, the number of proprietary medical schools had declined by almost half. By the early 1940s, out of the 155 medical schools that existed at the beginning of the century, only 77 remained (Angulo, 2017). This transformation was a significant step in improving the quality of medical education and ensuring that medical professionals were adequately trained and qualified.

Due to the absence of clear federal regulations, forming associations became a common way to address problems in higher education (Brittingham, 2009). Accrediting bodies have played a crucial role in ensuring the quality and standards of higher education institutions in the United States. Historically, accrediting bodies have been a cushion between the federal government and universities, although that distance has become smaller in recent years (Brittingham, 2009).

Relatively soon after accrediting organizations for professional schools were established, organizations developed to certify the quality of the growing number of colleges and universities in the United States. Organizations that fit the current definition of accrediting commissions emerged in the late 1800s (Brittingham, 2009). The New England Association of Schools and Colleges, established in 1885, focused primarily on preparatory and secondary school graduates in terms of their readiness for college. In the early years, accreditation concentrated heavily on the

specific question of whether an institution qualified as a college or university.

In the 1890s, there were reportedly over 900 higher education institutions, with the majority being relatively small and having an average of 160 students. However, in the mid- to late 1900s, the population attending college increased significantly, from 10% of 18- to 24-year-olds in 1945 to 30% by 1965 (Angulo, 2017).

As accrediting organizations began to be established, the New England Association of Schools and Colleges was one of the first regional accrediting bodies to arise. Other regional accrediting bodies followed suit, including the Middle States Association of Colleges and Schools, the Western Association of Schools and Colleges, the Southern Association of Colleges and Schools, and the Northwest Association of Colleges and Universities.

These accrediting bodies established specific criteria for becoming a member college or university, which can be traced back to the North Central Association in 1913. These criteria ensure that higher education institutions meet specific standards of academic quality, institutional governance, and financial stability.

In the mid-1900s, regional accrediting bodies began to develop the fundamental components of accreditation, including mission, standards, a self-study, a visit by a team of peers, and a report. Gradually, accreditors shifted their focus from input, such as teaching qualifications, to student outcomes. Factors such as graduate degree offerings, student-faculty ratio, and hours of teaching per semester were common dimensions in place by the 1940s. Other characteristics, although not necessarily mandatory for federal recognition, included the role of general education, academic freedom, and student development activities outside the immediate classroom, which were added by some accreditors (Angulo, 2017; Brittingham, 2009).

Currently, accreditation tends to take a future-oriented perspective with an emphasis on ideas for improvement. In the 1990s, universities were encouraged to adopt a Total Quality Management (TQM) approach, which was consistent with business trends (Sherr & Lozier, 1991). Under this approach, assessment results were regularly incorporated into the educational system for continuous improvement. TQM, initially designed for manufacturing settings, also has a customer orientation, which has become a visible presence in today's universities (Koris et al, 2015).

The Role of Government

Over the past 40 or 50 years, the involvement of Federal and State governments in higher education has increased significantly. However, the U.S. Constitution did not initially specify a role for the Federal Government in regulating education. In the United States, the responsibility for regulating higher education was traditionally left to the states, which mainly focused on protecting consumers. However, this approach made it challenging to assess the quality of higher education institutions, even as state and federal funding became the primary sources of financial support for universities (Thelin, 2011).

Governmental programs provide a significant portion of the funding for universities, with the Federal government being the primary funding source In the U.S., student loans and Pell Grants are significant sources of financial support for U.S. universities. Many institutions would likely not survive without this funding (Angulo, 2017). Universities must be accredited to receive funding, and accreditation is generally determined by non-governmental organizations in the U.S. and many European countries. It is essential to recognize that obtaining and maintaining accreditation can be critical to the success or failure of a college or university.

The Federal Government's involvement in accreditation was firmly established in 1952 with a revision of the original 1944 G.I. Bill, also known as The Servicemen's Readjustment Act (Baum et al., 2017). This updated bill mandated that veterans could only apply their education benefits to colleges and universities accredited by an agency recognized by the U.S. Department of Education.

In 1965, the U.S. Congress passed The Higher Education Act, which created Federal aid programs for non-veterans. Title IV of this Act reemphasized that only educational institutions accredited by organizations recognized by the U.S. Commissioner of Education could receive funds from federal sources such as Pell Grants and federally insured student loans.

The Higher Education Act was renewed in 1972. This renewed Act allowed Federal funding to be given directly to for-profit educational institutions such as vocational and business schools. The Act also extended government-backed student loans for use at for-profit institutions. Pell grants were also introduced in 1972. The only requirement for Federal assistance for someone attending a for-profit institution was that they had to be able to "benefit" from post-secondary education (Angulo, 2017).

As a result of these changes, the percentage of Federal assistance going to for-profit institutions increased dramatically. By 1982, 36% of Federally insured student loans went to for-profit institutions (Angulo, 2017). However, funding could not be directed to a higher education institution unless it was accredited.

In 1988, Secretary of Education William Bennett ordered all Federally accredited higher education institutions to provide data on institutional outcomes to accreditors. The 1992 amendments to the Higher Education Act provided more specific assessment requirements (Hannah, 1996). A later survey of university administrators' perceptions of assessment found that accreditors' requirements for documentation of student learning were the primary reason for conducting these evaluations (Marrs, 2009). In contrast, among those surveyed, issues like program or course improvement were not considered as important objectives of assessment (Marrs, 2009)

The 1992 Higher Education Reauthorization Act implemented specific standards for accrediting bodies to assess educational institutions (Hannah, 1996). These standards included criteria such as student achievement, curriculum, faculty, facilities, fiscal and administrative capacity, student support services, program lengths, student complaints, and compliance with Title IV program requirements. The Department of Education acknowledged five regional accrediting bodies at that time.

In 1998, the Higher Education Amendment focused on establishing quality standards for accrediting alternative education delivery methods, including online and various forms of distance education (Burd, 1998). The objective was to ensure Federal funds supported students in distance programs meeting acceptable quality standards. It was also emphasized that institutions that were not accredited should not present themselves to the public as accredited, as this was misleading to students and could have legal consequences.

The 2008 Higher Education Opportunity Act made important additions to accrediting practices, including requirements for due process in the event of an unfavorable decision by an accrediting body. It also mandated the need to publicly disclose policies on credit transfer between institutions. Moreover, accrediting agencies must have a process in place to evaluate student outcomes, particularly with respect to the institution's mission, state licensing exams, and job placement rates where applicable. However, it is important to note that these factors are not mandatory for the Department of Education's recognition of accreditation. Institutions and programs can

establish their own standards for student learning outcomes, and accrediting bodies would then assess the appropriateness of those standards (U. S. Department of Education, 2010).

Two national-level government organizations that oversee accreditation are the U.S. Department of Education (D.O.E.) and the Council for Higher Education Accreditation (CHEA). These organizations are not directly involved in the actual process of accreditation; instead, they provide resources, guidelines, and information about accreditation agencies. As noted earlier, some educational institutions publicly state that an agency accredits them. However, unless an accrediting body is recognized as an acceptable agency by the Department of Education, educational institutions are typically not allowed to receive federal funds. The Department of Education currently identifies the following functions of accreditors: 1) evaluating the quality of academic programs; 2) promoting continuous improvement; 3) involving faculty and staff in planning and assessment; and 4) for certain professions (such as nursing), establishing criteria for certification and, in some cases, licensure (U.S. Department of Education, 2024).

Currently, the following regional accrediting bodies are recognized by the U.S Council for Higher Education Accreditation:

- Accrediting Commission for Community and Junior Colleges (ACCJC) Western Association of Schools and Colleges
- Higher Learning Commission (HLC)
- Middle States Commission on Higher Education (MSCHE)
- New England Commission of Higher Education (NECHE)
- Northwest Commission on Colleges and Universities (NWCCU)
- Southern Association of Colleges and Schools Commission on Colleges (SACSCOC)
- WASC Senior College and University Commission (WSCUC)

During the Obama Administration, numerous reforms were advocated as part of a legislative package called "America's College Promise." Obama aimed to promote higher education for all Americans, with a particular emphasis on community colleges (Palmadessa & Palmadessa, 2020). A $50 billion increase in Pell Grants raised the student award to $5,775. Furthermore, Stafford loans, the government-guaranteed loan program for low-income college students, became available at reduced interest rates. President Obama also attempted to implement a comprehensive college scorecard. The scorecard's objective was to enable students and parents to

evaluate colleges and universities based on relative cost, debt incurred, employability, and time to degree completion. However, the scorecard was scaled down in response to lobbying from numerous higher education institutions. Consequently, while a large amount of data was made publicly accessible, navigating this information for practical decision-making regarding college selection proved challenging. The Obama Administration was particularly aggressive in addressing what was characterized as fraudulent and predatory actions by for-profit institutions (Angulo, 2017).

State Government and Higher Education

In the late 20th century, as higher education assessment became a profession, state legislatures began to require independent reviews of colleges and universities, separate from accrediting bodies. These mandates were in response to concerns about student learning, as expressed in *A Nation at Risk*, and were focused on consumer protection and economic outcomes, such as students' debt load (Cumming & Ewell, 2017). By 1990, around half of U.S. State governments required some form of higher education assessment, including learning outcomes.

Today, several states have passed or are considering laws requiring colleges to report information such as the number of graduates working in fields requiring a degree. As of 2023, Arizona, Arkansas, Kansas, Kentucky, Maine, Tennessee, Virginia, and West Virginia are among the states that require this information. Supporters of the current accreditation system argue that it promotes institutional self-regulation rather than relying on government standards. However, state governments are increasingly taking a regulatory approach and mandating that universities report graduation and employment rates. Critics argue that this regulatory approach may stifle innovation and diversity within higher education institutions.

Accrediting bodies have come under increased scrutiny in recent years, with concerns about their composition, the rigor of oversight, and the certification of for-profit institutions. Critics argue that the compliance-oriented nature of accreditation can discourage innovation in educational approaches and that certifications should be included in the accreditation process to focus on specific educational modules relevant to employment settings (Gaston, 2017).

Another issue raised by detractors is that accreditation does not provide a clear index of quality and tends to prioritize meeting standards over improvement. In addition, accrediting bodies have been criticized for not