Assessment and Accreditation in Business Schools

Robert S. Rubin and Kathryn Martell

Abstract

Business school mission statements routinely purport to promote some form of learning. Indeed, for most business schools, few issues are more mission-critical than the inculcation of competencies most important for the practice of management. Despite this significance, systematic efforts to demonstrate student acquisition of such knowledge and skill are not universally practiced. With increasing pressure for accountability from a multitude of stakeholders, business schools are being challenged to provide direct and systematic evidence of student learning. This chapter is intended to synthesize and augment knowledge of formal assessment practices within business schools and its role in the accreditation process. Specifically, current requirements and evidence to demonstrate student learning towards accreditation by the AACSB (Association to Advance Collegiate Schools of Business) are presented. In addition, the chapter details prevailing best-practices in assessment and discusses many of the perils often encountered along the way.

In the past few decades, higher education has been increasingly scrutinized by multiple stakeholders calling for improved accountability regarding the quality of education provided. Indeed, governments, students, accrediting bodies and the media to name a few are taking institutions to task for their inability to derive critical evidence indicative of educational achievement (Cabrera, Colbeck & Terenzini, 2001). Institutions have responded by attempting to develop measurements to capture the now well-known academic tripartite of research, service and teaching. Not surprisingly, of these major educational domains, institutions receive perhaps the most inquiries regarding their teaching and learning practices (Cabrera et al., 2001). Logically, external stakeholders peering into the ivory tower are concerned about what is in their view, the key role of the university: produce students who are more capable going out than they were coming in. Undeniably, stakeholders desire some assurance that their educational investment will provide the espoused returns institutions claim and are increasingly less willing to rely on vague promises of achievement and quality.
To date, however, many institutions have been rather reluctant to allow such external examinations of the sort being requested, choosing rather to rely on proxies (e.g., reputation, acceptance ratios, proportion of faculty with Ph.D.s etc.) as indicators of educational value. Such alternative measures are now being called into question. For example, in a controversial 2006 report commissioned by the United States Secretary of Education Margaret Spellings, the commission notes, ‘There is inadequate transparency and accountability for measuring institutional performance, which is more and more necessary to maintaining public trust in higher education’ (U.S. Department of Education, 2006: 13). Accordingly, the commission’s recommendations call for an overhaul of an institutional system that is primarily reputation-based to one that is performance-based. This would include, for instance, a ‘consumer-friendly information database ... to enable students, parents, policymakers and others to weigh and rank comparative institutions’ (p. 20).

Comprising such a system would be data obtained from national standardized college learning tests to allow for comparability across universities. True to the commission’s goals, major efforts are underway in the United States to create a national standardized college learning assessment (Kingsbury, 2007). Such trends and actions are creating considerable tension among institutional administrators and faculty who argue passionately for a continued system of independence from government oversight but also recognize the social and economic realities at play. In all, these trends are likely to continue and will have considerable impact on institutions’ ability to survive and thrive (Cabrera et al., 2001).

In response to the escalation for accountability, accrediting bodies are contemporaneously increasing pressure for universities to provide greater assurance of learning. Collegiate schools of business and their respective accrediting bodies have by no means been immune to these broader accountability trends. Perhaps no other institution has had a greater influence on the practice of learning assessment in business schools than the premier accrediting body, the Association for the Advancement of Collegiate Schools of Business (AACSB). The AACSB publishes a set of standards according to a minimum threshold concept. That is, they set out standards of which schools must demonstrate their minimal achievement. It does not establish a system of ranking or ‘aspirational’ standards, leaving schools to define their own standards of quality. Though these standards used to be quite broad, consistent with the broader trends, the most recent standards require institutions to complete assurance of learning measures that determine direct educational achievement (Thompson, 2004). Thus, institutions who state that they train certain managerial capabilities must present primary evidence that those skills have indeed been learned. As such, schools of business are being asked more and more to provide prima facie evidence of success.

With the above background in mind, the purpose of the present chapter is three-fold. First, we sought to briefly discuss the trends in management education leading to increased calls for learning assessment. Second, we provide an explication of the nature of accreditation standards and their associated requirements regarding assessment. Third, we sought to clarify methods for accomplishing assessment on programmatic level in schools of business, highlighting common traps along the way. In all, the chapter is an attempt to provide the reader with information that can be useful in understanding and applying learning assessment techniques in a dynamic institutional environment.

ACCOUNTABILITY IN SCHOOLS OF BUSINESS

Opportunities to systematically acquire managerial competencies are usually quite limited.
In most organizations, 70–90 percent of management learning occurs through informal work experiences, training, and mentoring (Pfeffer & Sutton, 2000; Tannenbaum, 1997). Although it should be noted that trial and error and informal experiences are effective, they tend to be somewhat inefficient means that require substantial investments, multiple years, and ability to synthesize unsystematic learning (Rubin, Baldwin & Bonner, 2002). With such limited access to systematic development, individuals hoping to acquire capabilities more efficiently are increasingly relying on university-based management education. The steady rise in graduate management education degrees awarded annually is quite reflective of this trend (Pfeffer & Fong, 2002).

The extraordinary growth in management education in the last few decades has had the added benefit for business schools of increased economic prosperity (Pfeffer & Fong, 2002). Indeed, as of the year 2000, 341 accredited schools of business offered graduate degrees in management representing well over 100,000 degrees awarded each year (Pfeffer & Fong, 2002). Needless to say, graduate management education is big business. Despite the rapid growth of business schools and the large numbers of newly minted MBAs produced yearly, the perceived value in the workplace of the MBA as a differentiating factor in managerial success remains high. In a recent survey (GMAC, 2006), corporate recruiters (n = 1,270) indicated that 40 percent of all new hires in 2005 held MBA degrees. In addition, these recruiters indicated that they would provide additional compensation and benefits to their MBA hires in the future and that starting salaries would continue to rise at a faster pace than non-MBA recruits. In all, individuals and organizations alike are increasingly turning to, and relying on, the business school marketplace for management skills and education. Such reliance on universities for broad managerial training is in part a mark of success in the long progression toward the professionalization of the management discipline (Porter & McKibbin, 1988).

Though business schools are enjoying unprecedented growth and economic prosperity, such ‘success’ has brought with it a whole new set of challenges. Fueled by evidence suggesting holding a graduate degree in management provides little extrinsic value (e.g., Pfeffer & Fong, 2002), a broad set of stakeholders, including governments, students, funding agencies, accrediting bodies, and organizations, are escalating their demand for increased accountability. Following simple logic, stakeholders purport that like any organization, investors ought to be able to determine some relative return on their investments. That is, management education ought to be viewed as a consumer-driven product (Zell, 2005). Regardless of the reason, however, the movement toward increased accountability in management education for what is being taught and learned is ubiquitous. Below, we briefly summarize the key accountability issues as emanating from three key business school stakeholders: students, organizations, and the media.

With this consumer-model in mind (Gross & Hogler, 2005), it comes as no surprise that student criticism toward business schools tends to focus on the methods used to develop future managers for the ‘real world.’ For example, Trank and Rynes (2003) noted that students are increasingly emphasizing employability as salient curriculum factors resulting in negative attitudes toward organizational behavior, human resources, management and anything that is not perceived as ‘useful’ in gaining employment. As such, students report decreasing tolerance for management theory (Trank & Rynes, 2003) and models. Following such logic, students routinely rank management courses at the bottom of the list of important courses in the business curriculum (Rynes & Trank, 1999). In addition, students' demand for curriculum that will land them jobs has impacted the evaluation system whereby faculty provide ‘easy’ courses in exchange for high course ratings. As a result, failures are rare and high rates of grade inflation exist (Pfeffer & Fong, 2002; Zell, 2001). Moreover, recent research
has shown that students exert considerable pressure on administrators to eliminate required management courses resulting in a curricula that is actually less relevant to managerial job requirements (Rubin & Dierdorff, in press). As such, students place significant pressures on faculty and administrators by asserting the criteria they believe to be most relevant in the assessment of success.

Magnifying the above concerns, organizations are exerting strong pressures on business schools to revamp their approach to training their future managers. Of particular salience is the move toward commodification of knowledge resulting in the training of tools and techniques that benefit organizations in the short-term (Trank & Rynes, 2003). That is, in order to meet organizations’ immediate needs in technology and specific job-focused arenas, organizations exert pressure on business schools to provide job candidates who can hit the ground running. As a result, business schools have increasingly turned to specialized programs and courses to provide job candidates who know how to use a particular tool, but may be less certain about the underlying complexities of the tool (Trank & Rynes, 2003). Perhaps the most visible pressure from organizations, however, emanates from corporate recruiters. Recruiters routinely assert that MBA programs need to do a better job at inculcating knowledge and skills as they relate to ‘soft-skills’ such as leadership, communication and interpersonal skills (Eberhardt, McGee & Moser, 1997; GMAC, 2006). Yet, research consistently demonstrates that recruiters select primarily on technical or ‘hard-skill’ competencies, to the exclusion of soft-skills.

Perhaps the most recent, influential and controversial stakeholder seeking to establish accountability is that of the media (Elsbach & Kramer, 1996; Gioia & Corley, 2002). In the form of business school rankings (e.g., Business Week, Financial Times etc.), the media has imposed its own set of standards that have to some extent virtually taken the place of any other indicators of success. Further, media rankings have become some of the most important variables that influence driving both student and recruiter choices (Safón, 2007). So critical are media rankings that they have been shown to factor heavily in administrative decisions such as curriculum design (Elsbach & Kramer, 1996) and retaining Deans (Fee, Hadlock & Pierce, 2005). From the perspective of the media, rankings are standardized, consumer-friendly and useful in helping prospective students, organizations and recruiters compare business school quality. As such, the media believe that such rankings make schools more accountable and establish a level playing field. In the process, however, Trank and Rynes note, ‘other than recruiter and student impressions, most other media measures of quality reflect either incoming student quality … or factors that vary by geographic locations … rather than measures of curriculum or teaching quality’ (p. 197).

Taken collectively, organizations, recruiters, students, faculty, the media, the government and others all seem to be laying claim to the criteria from which business school quality will be judged and ultimately will be held accountable. As noted above, pressures from these stakeholders for accountability are not simply a business school problem but follow the more general trends toward accountability. However, many disciplines such as medicine and law are a bit more shielded from such accountability pressures since their respective accreditation oversight (i.e., American Bar Association and American Medical Association) is highly standardized and drives program curriculum. Despite its role as the largest business school accreditation body, and unlike these other accrediting bodies, the AACSB is increasingly less inclined to enforce a unilateral approach to management education (Porter & McKibbin, 1988). Instead, the AACSB has been recently favoring flexible ‘mission-based’ standards that squarely places the onus of accountability at the institution level (Trank & Rynes, 2003). While the merits of such an approach can be debated, this approach does significantly impact the process by which institutions may attempt on the one hand to deal with
their multiple stakeholder pressures for accountability, and on the other hand to maintain their accreditation status. At the center of this tension are assessment standards and process. In the next section, we describe this process in depth.

**ACCREDITATION AND ASSURANCE OF LEARNING**

In the field of management education, the AACSB International and the European Quality Improvement System (EQUIS) are considered the premier accrediting bodies. Together, they have been responsible for the accreditation of over 600 schools of business worldwide. Although somewhat different labels for standards are often used, their approach to accreditation generally follows a comparable process (Urgel, 2007). For example, both bodies engage in a thorough review process in which institutions engage in self-study and teams of assessors perform an on-site visit evaluation of the school’s compliance with the standards. Further, most salient to the present chapter, both bodies require evidence that student learning has taken place. Because of these relative similarities, in the present chapter we chose to focus primarily on the larger of the two bodies, the AACSB.

Currently, 549 business schools hold AACSB accreditation, including 93 schools located in 30 countries outside of the US. AACSB schools comprise roughly one-third of business schools in the US, and 15 percent of business schools worldwide (AACSB, 2007). Once accredited, schools maintain their accreditation status every five years with required reports and peer review. The AACSB is routinely involved in public relations activities promoting the value of management education in general, and of AACSB accreditation in particular, to students, businesses, and the general public. The visibility and advocacy of the organization is an attempt to ensure accreditation adds value to the school’s reputation.

In order to achieve accreditation, business schools must satisfy a set of standards (the current version has 21 standards) related to the school’s strategic planning and resource allocation, faculty resources (including an evaluation of their research), and the quality and management of the curriculum. The most recent accreditation standards were passed by a membership vote in April 2003. The set of standards focusing on the quality of the curriculum are called ‘Assurance of Learning’ (standards 15–21). This choice of name reflects the emphasis on the assessment of student learning in attaining or maintaining accreditation. Prior to 2003, schools satisfied concerns over the quality of their curriculum by documenting their teaching and their curriculum review processes. Typically, schools would develop a matrix that would demonstrate how their business curriculum aligned with the topics and skill development that the AACSB considered mandatory. Many schools relied heavily on proxy data from students, alumni, and employers to further document the quality of their programs. Thus, ‘curriculum quality was evaluated primarily by inputs (e.g., Do we teach it?) and indirect outcome measures (e.g., student and alumni self-assessments: Do you think you learned it?)’ (Martell, 2007: 189).

This approach is no longer sufficient for AACSB accredited schools – schools must provide ‘hard evidence’ that their students are achieving the learning goals that form the basis for the curriculum. According to Milton Blood, former Director for Accreditation at AACSB, this is a significant change from the prior standards. In fact, he deemed it the most significant change between the 1990 standards and the 2003 standards remarking, ‘... the establishment of learning goals is going to be one of the greatest changes in how schools behave.... It is [an] evolutionary [change], but it is a very distinct change’ (Thompson, 2004: 429). While the reaction of those tasked with implementation was predictably negative, compliance with the AACSB AoL standards provides an opportunity to respond to critics of higher education discussed above. The new standards require business schools to respond directly by proving their students
are learning. A recent survey (Pringle & Michel, 2007) reveals that many business schools would not be expending considerable effort to document their students’ learning directly if not required by outside entities. Clearly, most business schools needed a push other than public criticism to provide direct evidence of the educational value of their degree programs. The AACSB clearly had this intention in mind, as evidenced by their language:

Few characteristics of the school will be as important to stakeholders as knowing the accomplishment levels of the school’s students when compared against the school’s learning goals... Assurance of learning to demonstrate accountability (such as in accreditation) is an important reason to assess learning accomplishments. Measures of learning can assure external constituents such as potential students, trustees, public officials, supporters, and accreditors, that the organization meets its goals. (AACSB, 2007: 60)

While the AoL standards did not figure prominently in the debate during the two-year review process, they have emerged as a key challenge for many schools seeking to maintain their accreditation. Very few schools had systems in place in 2003 that would meet the new requirements; furthermore, the assessment process set forth in the standards was so foreign that many schools were at a loss for where to begin:

As schools began to digest the new standards, a realization dawned that the AoL requirements were a major departure from what had been required in the past. The switch from indirect measures of learning (e.g., surveys) to direct measures (student demonstration of achievement) was significant. Moreover, most deans and faculty had no real idea what this evidentiary change even meant. (Martell, 2007: 189)

**Key AACSB AoL challenges**

The widespread reaction to the AACSB AoL requirements ranged from apprehension, to reluctance, to strong resistance. There are a number of challenges that must be overcome in order to make real progress in assessment: (1) Comprehension of the real purpose of assessment; 2 The change in focus from course grades to program assessment; and 3 The acknowledgement of the shift from indirect to direct measures.

A major challenge that must be overcome is helping faculty understand the true purpose of the assessment process. Some faculty are apprehensive that program assessment data will be used to evaluate their teaching or second-guess their grading (Pringle & Michel, 2007; Martell, 2007; Farmer, 1990), a response that has been characterized by some as ‘paranoia’ (Forgarty, 2003). Schools have systems in place as part of the evaluation, tenure, and promotion process – program assessment data should never be used for this purpose (Eder, 2005). Another issue related to perception of purpose is more subtle, but has the potential to reduce AoL to a bureaucratic exercise that adds little or no value to the curriculum. The purpose of AoL is not to prove to the AACSB that everything is going as it should, and that student learning is on track. Furthermore, it is not the intent to compare student learning in one school with that in another. The point of the AoL process is diagnostic; that is, to identify gaps in student learning and develop ways for which it can be improved in the degree program. This is the context and spirit in which AoL is intended to be approached, with an eye toward continual improvement.

A second challenge for many schools is the paradigm shift from strict course assessment to program assessment. Faculty routinely evaluate student learning of their subject material in the normal course of teaching. Usually a student’s grade in a course is a compilation of various learning assessments including exams, assignments, presentations, group projects, and so forth. Anecdotally, many faculty members’ response to the new AACSB assessment mandate was a variation of ‘I already do assessment, I assign grades.’ The focus of the AoL standards is learning across a curriculum, however; namely, a degree program:

Learning goals can be established at different levels in the educational process. At the course or
single-topic level, faculty members normally have very detailed learning goals. These standards do not focus on such detailed learning goals ... AACSB accreditation is directed at program-level learning goals of a more general nature. These goals will state the broad educational expectations for each degree program. These goals specify the intellectual and behavioral competencies a program is intended to instill. In defining these goals, the faculty members clarify how they intend for graduates to be different as a result of their completion of the program. (AACSB, 2007: 62)

Course grades reflect students' learning within a class, not across a curriculum. It is an important distinction that if unrecognized can stymie progress in achieving the AoL standards.

The classical assessment of Ph.D. students provides a good illustration of program assessment. The key learning assessments for Ph.D. students are written and oral comprehensive exams, a dissertation proposal defense, and the dissertation. These assessment methods provide the basis for faculty to evaluate the students' expert knowledge in their field, their ability to evaluate others' research, and their ability to design and implement a sound research project. Thus, the key focus in assessing candidates for Ph.D.s is not what they are learning in individual courses, but on what they learned as a result of their Ph.D. program. It is the faculty (plural) who develop the standards of performance and who conduct the assessments. Some specialized Master's level programs (e.g., MS in Finance) also often utilize comprehensive exams or projects that evaluate students' learning across their curriculum. But when we look at the degree programs that form the bulk of most business schools' portfolios – the undergraduate and MBA degrees – assessments of student learning are often exclusively conducted in the classroom for the sole purpose of evaluating learning in a single course. Prior to passage of the new accreditation standards, an overall GPA that met or exceeded a minimal requirement was often the only evidence that students had appropriately learned. Schools can no longer maintain AACSB accreditation with this approach.

Unlike individual assessment, program assessment calls for a broader explication of the programmatic learning goals. Faculty are now charged with thinking more broadly about their programs and to establish overall learning goals that all students, regardless of their major or what courses they select, should fulfill. Answering this type of question requires that faculty think outside of what happens in their classroom and across the curriculum. And if it is difficult for many to think across an entire curriculum when determining learning goals, it is even harder for many to imagine having to modify one's courses to address some of the deficiencies that AACSB AoL processes may uncover. For example, the most popular learning goal for undergraduate degree programs is 'effective communication' (Martell & Calderon, 2005). If assessment of students' writing skills finds them to be deficient, it is up to the business school faculty to develop a solution. This is true even if the faculty holds the conviction that writing should be taught elsewhere, identifying a skill as a learning goal means the faculty will be held accountable for students' performance.

Finally, a challenge facing many schools is providing the right form of evidence. The current standards call for a dismissal of indirect measures of learning such as focus groups and exit interviews. Accordingly, the AACSB maintains that such methods do not provide sufficient evidence of student learning. Gathering evidence on student learning through direct measures, however, is considerably more complex and administratively taxing than through the indirect measures which were in vogue prior to 2003. For most schools, these previous indirect methods of assessment were convenient and cost efficient and could be handled by administrative staff, whereas faculty play a pivotal role in collecting data through direct methods. Surveys conducted in the years following the passage of the new AACSB standards (Martell, 2005; Pringle & Michel, 2007) indicate that not knowing how to conduct direct assessments of program learning is a key source of faculty resistance to implementing
Building capacity for direct assessment practices

As recent surveys indicate, the current state or quality of AACSB AoL practices in schools of business is rather meager. A careful examination of these data reveals information about ubiquitous faculty resistance and fear of the process, apathy toward learning evaluation, and an apparent lack of perceived capability (Pringle & Michel, 2007). As one administrator recently remarked, 'a faculty member paid to do undergrad assessment didn't have a clue and wasn't respected by peers' (as cited in Pringle & Michel, 2007: 207). That assessment represents such a burden for institutions and their personnel alike is rather surprising given the professional training of most of its personnel, administrators, and faculty. Specifically, although the term 'assurance of learning' may be recently and inextricably tied to AACSB standards and accreditation, the process of evaluating student learning is by no means a new or convoluted endeavor. As noted above, the AACSB standards state, 'In defining these goals, the faculty members clarify how they intend for graduates to be different as a result of their completion of the program [emphasis added].' That is, how effective was the intervention of the degree program in bringing about its stated objectives?

In other words, the AACSB standards are at first more concerned with a summative rather than formative assessment. Formative assessment refers to the process of collecting information that can be used to improve the process of the intervention itself (Brown & Gerhardt, 2002). In more familiar terms, formative assessment examines both the process and content of what is taught in the curriculum. It is focused on the delivery of content and seeks to identify ways in which to improve it. Summative assessment, as the name implies, is focused on understanding whether or not the learning objectives set forth have been accomplished (Noe, 2006). It is true that the ultimate purpose of educational assessment is to generate data that are useful in improving the overall quality of the program; however, from an AACSB standards perspective, the place to start is with summative assessment that should ultimately inform a more focused formative assessment. Thus, following the AACSB, building an institution's assessment practices begins with the outcomes or effectiveness of the business degree program.

When framed this way, faculty and administrators should take some comfort in what it is that they are charged to do, namely, conduct applied research. Aside from the administrative burdens that are synonymous with documenting evidence, business school faculty and administrators readily have the tools to competently engage in assessment practices. Such tools were likely learned in their doctoral seminars and research meetings and are the same tools used to conduct their own research projects and evaluate individual student learning. Our view is that when institutions conduct assessment projects that generally meet good scientific standards of rigor and quality, they will de facto fulfill AACSB requirements.

Fortunately, the science of training and evaluation has much to offer in the domain of evaluating learning. Indeed, researchers have been actively seeking ways to target and measure learning in all types of occupational and educational settings for decades (Salas & Cannon-Bowers, 2001). Although disparate in nomenclature and steps, the extant literature suggest that comprehensive assessment of learning generally involves four critical practices, including: (1) Explication of learning objectives; (2) Development of evaluation criteria; (3) Selection of appropriate (matched) assessment methods; and (4) Decisions and use of data. In the paragraphs below we draw on this literature to address the key steps toward developing a successful assessment program that in the process will also fulfill AACSB requirements.
DEVELOPING STATE-OF-THE-ART ASSESSMENT PRACTICES

Learning goals and objectives

For purposes of this chapter, we take for granted that an institution has taken sufficient time to comprehensively conduct a needs analysis by establishing their mission, understanding who they serve and the critical knowledge, skills, and behaviors they hope to inculcate in their students at the program level. The outgrowth of this work leads to the first most critical practice in assurance of learning, establishing learning objectives. Learning objectives serve a number of important purposes. First, they help the school clearly delineate and communicate the most essential knowledge, skills, and behaviors that they hope their students will currently or ultimately possess (Quiñones & Tonidandel, 2003). Second, they establish the criteria from which all future evaluation will focus. As such, establishing good learning objectives should not be an afterthought, but is the foundation of good assessment practice.

The AACSB standards call for learning goals to be mission driven. To date, however, many business schools utilize mission statements that speak only in generalities about student learning. A survey conducted by the second author in 2004 (Martell & Calderon, 2005) found that the most popular learning goals for undergraduate business programs are knowledge about business concepts, communication, ethical reasoning, critical or analytical thinking, and teamwork. Further, Martell and Calderon (2005) also report that while 88 percent of AACSB accredited schools participating in the survey had identified learning goals, less than two-thirds had translated them into objectives. This is the second, critical stage in defining learning goals. Faculty must articulate what, exactly, they mean by concepts such as ‘critical thinking,’ ‘ethical reasoning,’ or ‘leadership.’ This process, which AACSB refers to as translating learning goals into learning objectives, must focus on student outcomes that are observable. For example, what could a student do, that faculty could observe, to indicate they possessed a global perspective? In debating this question, faculty may propose some of the following learning objectives: 1 Students will be able to identify the different aspects of national culture that affect work behaviors, and discuss how these cultural components impact motivation; 2 Students will be able to develop a marketing plan that is appropriate for a specified consumer good in an overseas market; 3 In a simulated setting, students will be able to exhibit culturally appropriate behaviors; and 4 Students will be able to correctly identify the impact of foreign exchange rate fluctuations on the income statement and balance sheet of a US company that manufactures and sells product in that country.

Transforming the above ideas into quality learning objectives usually requires attention to three key concepts (Noe, 2006). First, a quality learning objective is a statement of what the student is expected to know or do. Second, a specification of the level of quality to be demonstrated. Third, a clarification of the conditions under which a student is expected to demonstrate the knowledge, skill, or behavior, if necessary. This does not have to be a complicated or lofty statement. To the contrary, in order to apply across a degree program, most learning objectives are broad statements of student capabilities. Moreover, some have argued that there is often an overemphasis on writing highly detailed behaviorally-based learning objectives. As Kraiger (2002) remarked, one should, ‘feel free to write instructional objectives to reflect desired changes rather than a restrictive behavioral framework’ (p. 355). For example, the Kellstadt Graduate School of Business at DePaul University has established the following learning objective for communication in their undergraduate business degree program: Students will be able to communicate effectively both orally and in writing. They can produce a coherent written statement and oral presentation of the analysis of a complex business issue. Notice that both what the student is expected to know or do and the specification of quality (i.e., coherent),
though general in nature, are included. In many cases, the third component to learning objectives (i.e., conditions) will be rather straightforward (e.g., business environment). As programs increase in their specificity (e.g., Masters in Human Resources) the conditions may also increase in specificity (e.g., ‘in a unionized environment’).

Assessment criteria

Perhaps the most important and least understood aspect of assessment involves the specification of evaluation or assessment criteria. Important to remember, under the new AACSB standards, AoL is inherently learner-centered. In other words, although other target outcomes of assessment are important, such as institutional success (e.g., alumni giving, increased enrollment etc.) and educational delivery (e.g., number of courses, types of courses, etc.), the immediate goal of an assessment program is to focus on changes in the learners to demonstrate program efficacy. Although many models have been proffered to guide the explication of target outcomes (e.g., Kirkpatrick, 1976), one recent model developed by Kraiger, Ford and Salas (1993) has served to significantly augment understanding of learner-centered criteria. Importantly, this model draws its value from greater than 50 years of learning theory (e.g., Anderson & Krathwohl, 2001; Bloom, 1956; Krathwohl, Bloom & Masia, 1964; Gagne, 1984). Thus, the most significant contribution of the Kraiger et al. (1993) model is the extensive synthesis of prior research (e.g., Bloom’s Taxonomy) resulting in a comprehensive yet parsimonious model of the most time-tested learning taxonomies and classification systems. Specifically, Kraiger et al. outline three overarching learning outcomes and their related learning concepts: 1 Cognitive outcomes; 2 Skill-based outcomes; and 3 Affective outcomes. Each of these outcome categories is discussed in more detail below.

Cognitive learning outcomes. Cognitive outcomes refer to the acquisition of knowledge, both in terms of the quantity and the type of knowledge learned. In addition, cognitive outcomes represent the degree to which learners demonstrate the relationships among particular knowledge components (Kraiger et al., 1993). As such, cognitive learning outcomes can be categorized into three major domains: 1 Verbal knowledge; 2 Knowledge organization; and 3 Cognitive strategies. Verbal knowledge, often referred to as declarative knowledge, is the traditional focus of university-based assessment and involves assessing the amount and/or accuracy of knowledge acquired. Most college instructors’ exams are aimed at measuring students’ declarative knowledge. Knowledge organization refers to the way in which students organize or categorize their knowledge. This differs from declarative knowledge in that it goes beyond simply possessing the knowledge to include an understanding of the interrelationships between knowledge structures. Finally, cognitive strategies involve learners’ capabilities to select the most appropriate form of knowledge to apply to learning, remembering, and problem solving (Gagne, 1984). Often referred to as metacognitive skills (Brown, 1975), this form of cognitive outcome involves planning and revising one’s actions based upon previous knowledge of the problem or situation. Although the reader may be less familiar with the specific aspects of cognitive outcomes, they are perhaps the most commonly and easily captured in an academic setting as we discuss in the section under assessment methods.

Skill-based learning outcomes. Sometimes called behavior outcomes, skill-based learning outcomes involve the demonstration of technical or motor skills that learners had not previously held or demonstrated as well as the capacities to perform these skills with fluidity under real conditions (Kraiger, 2002). Learning objectives focused on students actually ‘doing’ are generally attempting to capture some form of skill-based outcome. Two forms of skill-based outcomes are important to note. First, skill-acquisition involves both proceduralization (i.e., the ability to perform newly
acquired behaviors) and compilation (i.e., the increasing capability to demonstrate behavior without error). For example, as learners acquire skills in programming a spreadsheet, they learn to apply various steps in the process (proceduralization). With practice, learners begin to move away from the separate steps involved and treat the entire process as a whole (compilation). Second, ultimately learners’ skill leads to automaticity whereby the skill learned requires increasingly less effort thereby freeing capacity for other tasks. Although universities routinely purport to measure skills, few institutions focus their assessment efforts on these types of outcomes (Bommer, Rubin, Bartels, 2005).

Affective learning outcomes. Affective outcomes refer to learners’ attitudes or motivation toward the particular learning objective. These outcomes include understanding the direction or strength of particular attitudes such as a learners’ conviction. In addition, the motivation aspect of affective outcomes includes learners’ confidence levels, also known as self-efficacy. Such outcomes, sometimes called ‘reaction measures’ (Kirkpatrick, 1976) are usually quickly dismissed by those who seek ‘hard evidence.’ However, significant research indicates that attitudes are related to subsequent performance and skill acquisition. For example, there is strong evidence linking individuals’ levels of task-specific self-efficacy to subsequent attempts and performance of the task (Martocchio & Hertenstein, 2003; Stajkovic & Luthans, 1998). Further, affective outcomes have been narrowly defined in previous research (Kraiger et al., 1993) and thus are often seen as inconsequential earning terms such as ‘smile sheets.’ But as Gagne (1984) notes, ‘schools do a great deal to establish attitudes. Schools are fairly successful in establishing socially beneficial attitudes’ (p. 383). Indeed, with the heightened emphasis on ethics, corporate social responsibility, and citizenship, many institutions are in fact highly concerned with the attitudes and motivations of their students. Measuring affective outcomes in ways that are useful and valid can be difficult. For this reason, affective outcomes are often discouraged by AACSB examiners in favor of skill and cognitive outcomes. This is unfortunate, as much can be learned from rigorously designed affective assessment outcomes.

Problems in criteria development. In selecting criteria that will sufficiently capture the complexity of the established learning objectives, it is important to be aware of some common problems associated with criteria development. One area of particular importance is that of criteria deficiency and contamination (Noe, 2006). Criteria deficiency is present when the specified criteria only measure part of what is hoped to be learned and expressed by the learning objective. For example, consider the following learning objective, ‘students will be able to use analytical and decision-making skills in problem solving.’ Clearly this objective involves cognitive and skill-based criteria. To capture only cognitive criteria for instance would be a rather incomplete assessment of this learning objective. Conversely, criteria contamination involves measuring criteria that are not reflected in the desired learning outcomes or criteria that are impacted by outside factors. Using the above example, the learning objective does not seek to measure issues surrounding affective outcomes such as motivation or attitudes toward decision making or problem solving. Measuring learners’ attitudes toward problem-solving, while valid in its own right, would introduce criteria not called for by the learning objective itself. Problems of contamination and deficiency can be avoided by ensuring that the criteria chosen are the most relevant to measuring the stated learning objectives.

Assessment methods

As we noted above, in order to truly assess learning, one must specify the actual knowledge, skill, or behaviors associated with a given learning objective. This process entails identifying what will be measured (e.g., non-verbal communication) and also how it ultimately gets measured (e.g., a classroom presentation). How the learning goal gets measured is a process we refer to
as selecting appropriate assessment methods and should not be confused with the establishing criteria (i.e., what gets measured). Assessment methods are neutral; that is, they are a mechanism by which a particular capability is elicited and captured. Further, many methods can and should be used effectively to capture more than one capability. For example, an institution hoping to measure decision making may use a written test to measure knowledge of the specific target (e.g., stakeholder analysis) and a case study to elicit proceduralization of the decision-making steps (e.g., determining the appropriate stakeholders for a technological change effort). In addition, since the learner is providing a sample of writing, written communication might be assessed as well.

Selecting appropriate assessment methods can be confusing. The most important factor to consider in choosing the right method is the specific learning objective that is being targeted. The goal is to match the objective capability with the assessment that will best elicit and measure that capability (Quiñones & Tonidandel, 2003). For example, institutions routinely target interpersonal skills such as leadership and teamwork in their learning objectives, yet when it comes time to assess these skills, they employ methods more suitable to cognitive outcomes such as written exams (Bartels, Bommer & Rubin, 2000). That is not to say that a body of knowledge in the realm of leadership or teamwork is irrelevant; there is of course an important and extensive knowledge base to master. Yet, if the goal is to develop individuals who can influence people effectively, a method such as role-play or work-sample (e.g., vision speech) provides a much more salient context for which to display these skills.

Another important aspect of choosing the right method is to focus on the most appropriate source of evidence to support that learning has taken place. If the learning objective is targeting awareness of ethical issues, for example, the most appropriate data source would be the learner (e.g., self-report survey). However, if the learning objective is more concerned with communicating an ethical message, the data source may be ratings from an instructor (e.g., rubric of communicating ethical ideas).

The process of selecting an appropriate method can be simplified by relying on classification schemes. The classification scheme by Kraiger et al. (1993) later elaborated by Kraiger (2002), can easily be modified to suit the university learning environment and can serve as a road map for selecting the most appropriate method. In Table 19.1 we provide a summary of assessment criteria and potential methods most appropriate for assessing the criteria. As mentioned above, some methods are more conducive than others for capturing certain types of criteria. The best approaches use multiple methods in an effort to 'triangulate' the learning objective. Further, Table 19.1 should be viewed as illustrative rather than exhaustive since skilled assessors and well-crafted assessments can often elicit and capture multiple forms of outcomes. Below, we briefly describe some advantages and disadvantages common to assessment methods.

**Measuring cognitive outcomes.** The quintessential cognitive assessment is the written examination or 'paper-and-pencil' test. Traditional methods familiar to the reader such as multiple-choice, true-false and so on are best suited for learning objectives which require a demonstration of accurate recall or recognition. Working knowledge of most disciplines is captured quite well by these tests. In attempting to measure learners’ cognitive strategies, however, it is more difficult to create multiple-choice test to appropriately elicit such criteria. Appropriate methods might include problem sets whereby students can show their work or case scenarios where students can explicate their higher-order thinking to solve a case (for an example of more sophisticated methods see Davis, Curtis & Tschetter, 2003). It should be noted, however, that research suggests that knowledge examinations are best administered close in time to the point of learning (Kraiger et al., 1993).
Table 19.1 Classification of assessment learning outcomes

<table>
<thead>
<tr>
<th>Learning outcome category</th>
<th>Learning concept</th>
<th>Measurement focus</th>
<th>Potential assessment methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive outcomes</td>
<td>Verbal (declarative) knowledge</td>
<td>Quantity of knowledge, recall accuracy, speed of recall</td>
<td>Exams testing recognition (e.g., multiple-choice) or recall (e.g., essay, fill-in-the-blank), Concept mapping or card sorting</td>
</tr>
<tr>
<td></td>
<td>Knowledge organization</td>
<td>Idea similarity, knowledge interrelatedness, hierarchical ordering</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cognitive strategies</td>
<td>Forming concepts and procedures, problem solving, Proceduralization compilation</td>
<td>Case scenarios, problem sets</td>
</tr>
<tr>
<td>Skill outcomes</td>
<td>Skill acquisition</td>
<td></td>
<td>Assessment centers, work samples, role plays, behavioral checklists, presentations, Behavioral observation, performance ratings, Self-report, task specific self-efficacy</td>
</tr>
<tr>
<td></td>
<td>Automaticity</td>
<td>Automatic processing</td>
<td></td>
</tr>
<tr>
<td>Affective outcomes</td>
<td>Attitude</td>
<td>Targeted object (e.g., ethics), attitude strength, self-efficacy</td>
<td>Self-report, observation, time-on-task, goal-difficulty ratings</td>
</tr>
<tr>
<td></td>
<td>Motivation</td>
<td>Effort, tenacity, goal difficulty, motivation to learn</td>
<td></td>
</tr>
</tbody>
</table>

Source: Adapted from Kraiger (2003) and Kraiger et al. (1993)

Measuring skill-based outcomes. Assessing skill-based outcomes usually involves a more active form of testing that includes observing student behavior and utilizing some form of standardized rating system from which to form judgments about skill capabilities. Skill-based measures might include work samples, role plays, behavioral checklists, presentations, or even observation. Whether judging student proceduralization or some deeper internalization, skill-based outcomes most surely involve collecting multiple observations. The reader is likely familiar with the use of role plays and oral presentations in assessing skills. Two methods less familiar, but highly effective are work samples and assessment centers. As the name implies, a work sample is a sample of an actual piece of a job (Guion, 1998). Used to select employees, work samples are designed to test discrete job behaviors through the demonstration of particular skills in a realistic environment. Work samples are among the best predictors of future work performance, as those who can demonstrate the skill today will likely be able to demonstrate it in the future (Roth, Bobko & McFarland, 2005; Schmidt & Hunter, 1998). As such, translating this process to academics holds promise for effective assessment. For example, using work sample technology to assess students' specific vocational skills in technical areas such as marketing presentations, financial analyses, or performance appraisals would allow for broad assessment of a set of important skill-based outcomes.

Another emerging technology in the use of skill-based outcome measurement is that of the assessment center. Used for decades as an employee selection tool, the assessment center can also be used to collect systematic information about student skills (Waldman & Korbar, 2003), particularly interpersonal or 'soft' skills (Bommer et al., 2005). Assessment centers place students in a hypothetical scenario where they typically adopt the role of a manager or employee in a fictitious company. Through interactive exercises such as meetings and oral presentations designed to elicit specific skills, trained assessors evaluate student skill domains such as initiative, leadership, planning and organizing and decision making. Since performance in student assessment centers is linked to important career outcomes (Waldman & Korbar, 2003), performance results can be
fed back to learners in a meaningful way for skill development (Kottek & Schultz, 1997; Mullins, Shaffer & Grelle, 1991). Although assessment centers hold a great deal of promise for assessing certain skills, they do come with considerable administrative barriers, such as costs, faculty time, and expertise. Hence, to date their prevalence as an assessment technique remains rather low (Bommer et al., 2005).

Like cognitive outcomes, unless skills are truly practiced throughout a program’s curricula, measuring skills close in time to when they were learned is vital to gain an accurate picture of the skill. In fact, meta-analytic research demonstrates that after one day of training little to no skill decay exist. Yet, one year post-training results show a 92 percent decrement in skill demonstration (Arthur, Bennett, Stanush & McNelly, 1998). This research highlights the potential problems inherent in the so-called ‘end of program’ exam and comprehensive field exams. Students who do not continue to utilize the knowledge and skills taught early in the program will not likely demonstrate those skills later in their program. This may in part explain the trend away from field tests toward course-embedded assessment (Pringle & Michel, 2007).

Measuring affective outcomes. Since affective criteria exist solely within the person, the demonstration of change in affective criteria is most typically measured by self-report instruments. As noted above, the most common form of affective outcome measures are reaction outcomes whereby learners indicate their satisfaction with the content and process of a particular exercise, course, and/or program. Such measures are routinely captured by universities post-term to gauge student experiences with the course. Although useful toward understanding what learners believed to be valuable in enhancing learning, such measures provide little evidence of actual learning. Indeed, reaction measures have received strong criticism in the assessment literature (Holton, 1996) for their lack of association with actual learning (Alliger & Janak, 1989).

Recent developments in affective criteria, however, have given way to more viable forms of affective outcome measurement. One particularly useful affective outcome is that of self-efficacy, generally defined as a learner’s belief in his/her ability to accomplish a given task (Bandura, 1977). Self-efficacy has been conceptualized to be a vital motivational factor that directly influences how individuals approach challenges in different learning situations. Empirical research has supported this notion showing positive relationships between self-efficacy beliefs and learner task choice, effort, and persistence in overcoming challenges (Bandura, 1991; Gist & Mitchell, 1992), training performance (Tannenbaum, Mathieu, Salas & Cannon-Bowers, 1991), and reactions to training (Gist, Schwoerer & Rosen, 1989). Within the learning environment, it is often the case that objectives directly seek to increase learner confidence in a particular task (Kraiger et al., 1993) such as public speaking or business writing. Thus, given the extant research, assessment programs may benefit greatly from the inclusion of self-efficacy measures as a gauge of learning.

Another important affective outcome to consider is that of motivation to learn which has been defined as the desire to learn content within a given program or developmental environment (Noe, 1986). Motivation to learn is thought to influence learners’ decision-making processes critical to the effectiveness of the developmental program and is positively related to learning (Colquitt, LePine & Noe, 2000), participation in developmental activities (Noe & Wilk, 1993), positive reactions (Cole, Field & Harris, 2004), and completion of programs (Baldwin, Magjuka & Loher, 1991). In addition, learning motivation has been shown to be associated with increased transfer (actual use of the acquired knowledge, skills, or behaviors) to work environments (Colquitt et al., 2000). In light of these findings, it is important to note that learning motivation is generally seen as a malleable attitude (Cole et al., 2004). Thus, efforts to assess change in learning motivation can provide meaningful
information as part of a comprehensive assessment program.

Although we have discussed multiple methods for assessing learning outcomes, we by no means have provided an exhaustive review. As we noted, the 'best' method will always depend on the specific learning goal and the method(s) most suitable for measuring it. Further, we hasten the reader to note that assessment methods, regardless of how sophisticated or elaborate, contain imperfections. As with any evaluative domain, errors associated with the assessor and deficiencies in the instrument can skew results. Concerns about adequate reliability (i.e., consistency in measurement) and appropriate discriminability (i.e., that the measure reflects actual differences in the criteria) are always important considerations in selecting or developing an assessment method (Noe, 2006).

**Assessment designs**

Aside from determining the appropriate criteria and the methods that best measure the criteria, the design of assessment efforts plays an important role. The design of an assessment effort refers to a plan for the circumstances under which assessment information will be elicited and collected. The reader is no doubt familiar with the various evaluative design paradigms available, such as post-tests, single group pre-test-post-tests, pre-test-post-tests with comparison groups and so forth. Each of these designs is associated with various costs and benefits as they relate to controlling threats to the internal and external validity of assessment results. A full discussion of these designs is beyond the scope of this chapter and is discussed extensively elsewhere (see for example Cook & Campbell, 1979; Cook, Campbell & Peracchio, 1990). Thus, rather than provide full treatment here, we instead discuss some recent alternatives to the classic quasi-experimental designs. First, however, we discuss designs put forth by the AACSB under the new standards.

It should be noted that the language we use in this chapter is not entirely consistent with the AACSB's documentation. This is intentional as we aimed to be more precise in our language. In particular, the AACSB refers to the designs below as 'assessment approaches.' Unfortunately, this language obscures the distinctions we have made between matching criteria to the appropriate method (methods such as written exams for cognitive-based outcomes) and design choice (structure of the research). Nevertheless, because of their criticality toward accreditation, we briefly discuss the AACSB AoL standards with respect to assessment design. Three assessment designs have been offered by the AACSB as ways to collect assessment data: 1 Selection; 2 Course embedded; and 3 Determination. These are briefly discussed below followed by a section devoted to increasing assessment design rigor.

**AACSB assessment designs.** Selection designs refer to an assessment that takes place at the point of entry. For example, to gain admission to the Neely School of Business at Texas Christian University (Wakefield, 2005), students must demonstrate their proficiency with business software packages. Another example is provided by the C.T. Bauer College of Business at the University of Houston (Anderson-Fletcher, 2005), which requires applicants to demonstrate a minimal level of writing competency in order to be admitted. The selection design is not frequently used for AACSB AoL purposes (Martell, 2007; Martell & Calderon, 2005), and is only a viable option for undergraduate business programs that admit students in their sophomore or junior years.

The second design approach listed in the AACSB documentation, course-embedded, is the most widely-used design in business schools. In recent surveys (Martell, 2007; Pringle & Michel, 2007) 70 percent of the respondents indicated they used the course-embedded design. In the course-embedded design, products generated by students to fulfill requirements for a course are also used for AACSB AoL purposes. The course-embedded design collects student performance data across multiple courses using an established performance standard.
For example, a course product (e.g., a case write-up) is assessed by one professor for his/her course using his/her own performance criteria for the purposes of assigning a grade, while a second copy of the paper is assessed by someone else using the criteria determined by the faculty at large. A comparison of the outside evaluation against the professor’s grade is not part of the assessment process. In other words, the course serves as the mechanism by which student assessment data are elicited. Course embedded methods do have many practical advantages. First, they take advantage of student products already incorporated in the curriculum which simplifies the assessment process. Second, they tend to be closely tied to the school’s curricula. Third, by utilizing student work that impacts their course grades, students’ motivation tends to be high thereby capturing students’ best efforts.

Demonstrations or tests are the final direct assessment design discussed in the AACSB AoL standards. These assessments take place outside of the classroom, in spirit if not in actuality. Often referred to as the single-group post-test design, demonstrations are rather straightforward – a single group (e.g., senior business students) completes a test or battery of tests designed to evaluate multiple learning objectives. As we noted above, Doctoral programs rely on demonstration designs when evaluating candidates via qualifying or comprehensive exams. In undergraduate assessment programs, the most commonly used method within a demonstration design is the Educational Testing Service (ETS) Major Field Achievement Test in Business (MFAT-B). More than 500 business schools use the MFAT-B to evaluate their students’ cognitive-based outcomes.

Methods such as the assessment center used by the Kellogg Graduate School of Business at DePaul University, the Stillman School of Business at Seton Hall’s assessment panel (Boroff et al., 2005), and the mock interviews used at the School of Business at Montclair State are all methods that lend themselves nicely to the demonstration (i.e., single group post-test) design. Demonstration designs have some key advantages, namely, they are comparatively easy to set up and, depending on the method used within, can assess students on multiple learning goals simultaneously. For example, the assessment center used at DePaul University assesses oral communication, teamwork, and decision making in the same exercise, and the Stillman panel evaluates oral communication, knowledge, teamwork, and critical thinking. Potential drawbacks include costs (the MFAT-B is more than $25 a student, developing an assessment center is very time intensive, the Stillman School’s use of alumni as outside assessors is a time-consuming administrative task), lack of alignment between the school’s curricula and measures developed elsewhere, and student motivation. Once assessment activities are taken outside of the classroom, students’ motivation to provide their best performance may be diminished. Even if the school makes participation mandatory, the issue of motivating students to give their best effort must be addressed by other means. A typical response is to incorporate students’ performance on the demonstration into a course grade (see Rotondo, 2005 for a discussion of other motivating techniques).

Increasing assessment design rigor

Although we encourage institutions to think in terms of increasing the rigor associated with their assessment designs, such rigor does not necessarily require multi-group pre-post quasi-experimental designs. Indeed, when one reflects on the considerable costs, feasibility, and availability of resources to conduct such assessment, the process can be rather daunting. Thus, institutions should determine whether or not they actually need to demonstrate change in the learner or whether achievement of a particular standard is sufficient. For many institutions, their concern is that students achieve a particular level of competence in relation to some standard, rather than improvement or that future student learning will be improved.
In such a situation, a single group post-test (course embedded or determination design) can suffice. However, it is important to keep in mind that such a design does not evaluate the efficacy of the program in inculcating the criteria it espouses, rather it evaluates the level of knowledge, skills, or behaviors in its students, regardless of how or where it was acquired.

Institutions interested in showing changes in the learner can employ some alternative approaches which can improve rigor but do not require a classic pre-post test control group design. Haccoun and Hamtiaux (1994) introduce an alternative design called the internal referencing strategy. This design allows for increased causal inferences as the authors describe:

The internal referencing strategy may be described as a pre-post single group design in which the evaluator purposefully incorporates in the pre-and post measures items which are relevant to the training (i.e., which ought to change because the course content will cover them) and items which are not expected to change (because the course will not deal with them). Comparisons are then established between pre-post differences on the relevant items as well as on the irrelevant items. Effectiveness is inferred when changes on relevant items are greater than changes on irrelevant ones. (p. 595)

As might be apparent, this strategy applies most readily to assessment of cognitive-based outcomes, though as Haccoun and Hamtiaux (1994) note it may with some effort be applied to behaviorally-based outcomes.

Another alternative design to consider in lieu of an actual control group (i.e., who never receives the educational intervention) is to adapt the staggered training or rolling group design (Kraiger, McLinden & Casper, 2004; Quiñones & Tonindandel, 2003) to educational environments. In this design, training is given to all participants; however, training groups start at different points in time. Groups waiting to begin training can serve as the comparison group to groups who are already in progress. Educational environments provide natural rolling groups, particularly for course embedded or program assessment. For example, a group that completes post-test measures at the end of a term as part of a business writing course could be compared against a group’s pre-test measure at the beginning of the term.

‘Use of assessment data for program improvement

The AACSB, in addition to other accrediting bodies and assessment scholars, are clear about the ultimate purpose of assessment: to assure student learning.

Measures of learning have little value in and of themselves. They should make a difference in the operations of the school. Schools should show how Aol. results impact the life of the school. Such demonstrations can include uses to inform and motivate individual students and uses to generate changes in curricula, pedagogy, and teaching and learning materials. (AACSB, 2007: 69)

From this perspective, analyses should focus on what has been learned about students’ learning. In doing so, areas for curriculum improvement will most certainly be identified. This process, which is referred to by the AACSB as ‘closing the loop,’ is one of the greatest concerns that business school Deans have about assessment (Martell, 2007). Utilizing assessment results is critical because the AACSB requires ‘the impact of assessment outcomes on continuing development of the degree programs ... be evident’ (AACSB, 2007: 69).

In other words, it is not enough to examine a single outcome and implement a change as a result of the examination; rather, continual systematic efforts must be made to develop the degree program until evidence can be shown that the intervention(s) has achieved its goal. More complex is that the summative evaluation often gives way to formative evaluation as schools seek answers to why there are particular deficiencies in knowledge, skills, or behaviors. In such a process, institutions will likely explore issues related to andragogical techniques, course integration, as well as admission procedures and the rigor of the program.
Matching the appropriate method to criteria as discussed previously allows for a more parsimonious search for future interventions. For example, if a particular cognitive-based outcome such as knowledge of employment law is relatively low among students, the interventions proposed will logically follow this finding (i.e., interventions design to improve learning and retention of employment law). The improvement process is a cyclical one much like those championed by the total quality management movement. Here the goal is to make data-driven decisions toward improvement, track the data and feed it back into the decision-making system. In this spirit, many institutions are adopting a two-year cycle in which progress on learning goals are routinely and systematically assessed and development interventions implemented.

CONCLUDING THOUGHTS

Throughout this chapter, we have attempted to illuminate the trends associated with assessment and clarify the requirements for accomplishing assurance of learning toward accreditation. We have of course omitted a great deal of information related to assessment concerns, such as andragogical approaches and their impact on student learning, as well as practical and logistical recommendations for accomplishing assessment. For example, recent research has shown that capturing individual difference variables such as grade point average (Bacon & Bean, 2006) alongside other assessment information can allow for increased understanding of the outcomes of interest or effect of a particular intervention. In addition, we did not provide a comprehensive list of tools or techniques to accomplish assessment. Although important, we hoped that this chapter would provide a more basic foundation for the accomplishment of assurance of learning. Indeed, it has been our experience that institutions, in their haste to establish an assessment process, often benchmark other universities and blindly adopt the assessment tools they uncover. Rightfully so, off-the-shelf tests and techniques tested elsewhere are seductive and seemingly come with turn-key solutions to assessment. Unfortunately, not unlike corporations who rely heavily on benchmarking, institutions who adopt methods without a thorough consideration of criteria are not likely to effect real change in their programs (Pfeffer & Sutton, 2006). We hope that the information presented here assists schools in employing a more thoughtful and systematic approach to their assessment programs.

NOTE

1 For more extensive reviews of assessment centers in academia see Bommer et al. (2005) and Riggio, Mayes, and Schleicher (2003).

REFERENCES


Rubin, R. S. & Dierdorff, E. C. (in press). How relevant is the MBA? Assessing the alignment of required