Assessment of Student Learning in Business Schools: Best Practices Each Step of the Way

Edited by
Kathryn Martell, Montclair State University
Thomas Calderon, University of Akron

THE ASSOCIATION FOR INSTITUTIONAL RESEARCH
and AACSB INTERNATIONAL
Volume 1; Number 1
Assessment in the Disciplines

Table of Contents

FOREWORD
Jerry E. Trapnell
Executive Vice President and Chief Accreditation Officer
AACSB International

For almost 90 years, AACSB accreditation has represented the highest standard of quality business education. As expectations for higher education have evolved, so have our accreditation standards. In the early 1990s, AACSB switched from accrediting business schools based on the strength of their resources and reputation to a new mission-based philosophy. In order to earn accreditation, business schools must clearly establish a mission and corresponding strategic goals, and demonstrate progress and continuous improvement relative to those goals. While a business school may have many different goals, none is as fundamental as student learning.

In 2003, the AACSB membership approved a new set of standards that places student learning at the forefront of accreditation. In doing so, our members explicitly recognize student learning as the most fundamental goal of the academy, and that we must document to our stakeholders that graduates of accredited business programs will have achieved specific educational outcomes. The new Assurance of Learning (AOL) standards— which require that business schools articulate their learning goals for their students, and demonstrate progress and continuous improvement in curriculum relative to those goals—will help all of us to maintain that focus.

Consistent with our mission-based approach to accreditation, AACSB recognizes that business schools will pursue different learning goals that depend on the nature of their institutions and the context in which they operate. Thus, it is not expected that different business schools will pursue a homogenous set of goals. At the same time, AACSB expects that schools will specify their mission-appropriate learning goals and will "demonstrate achievement of learning goals for key general, management-specific, and/ or appropriate discipline-specific knowledge and skills" (AACSB, 2004, p. 16). While schools are given considerable latitude in the choice of goals and assessment methodology, they must provide **direct** outcome evidence of students' progress in meeting learning goals, and modify their degree programs in light of their findings. This is a departure from the past, when indirect (survey) measures provided the preponderance of evidence on the effectiveness of business schools' curricula.

While AACSB has taken a leadership position in establishing student learning standards, we are not alone. As demands for accountability from stakeholders including the business community, governing boards both public and private, and federal and state legislative bodies grow louder, regional and professional accrediting agencies are modifying their requirements to include student learning outcome measures. Although many schools may initially view the AOL standards as a challenge, they also offer us an opportunity to respond to our critics who question the value of business education.

In this period of transition to the new standards, there is a hunger for information. The examples generously provided by the authors of chapters in these volumes provide ideas, advice, and encouragement to help schools establish their own meaningful assessment programs. On behalf of AACSB-International, I would like to thank the editors, Kathryn Martell and Thomas Calderon, for a significant contribution to the assessment literature. I would also like to thank the Association of Institutional Research for their sponsorship of this project. Assessment of student learning is as much about the conversation as it is about the outcomes. These books give us a lot to talk about.

AACSB. (2004). Eligibility procedures and standards for business accreditation. Tampa, FL: AACSB

FOREWARD

In the fall of 2004, Dr. Richard Howard of University of Minnesota asked me, on behalf of the Association for Institutional Research (AIR) Publications Committee, to consider editing a series on assessment in various academic disciplines. The interest in such a series has grown out of a concern that assessment in many academic departments and fields of study is a somewhat neglected area and is quite variable across different disciplines. Most of the publications regarding assessment which do exist, if they discuss disciplinary assessment at all, group the information in such a way as to make it difficult to select effective methods in a single field of study. Even the resources on the Web site maintained by Dr. Ephraim Schechter at North Carolina State University (http://www2.acs.ncsu.edu/UPA/assmt/resource.htm), one of the best single sources on assessment on the planet, are listed primarily by institution and not by discipline, because that is the way that the information is made available.

Another factor making a volume focused on a single field of study so desirable is the simple fact that different disciplines have different cultures and vocabularies and even different pecking orders. This is very apparent to one such as myself who works at a large and diverse university with many departments; a different language is spoken in every department, and sometimes even those in sub-disciplines within a department do not understand each other. It's like going to a foreign country and encountering not only the official language but also various regional dialects. Too many departmental visits in one day can lead to a form of culture shock.

This volume and the ones to follow it are the first in what we hope to be a long and varied series on assessment in the disciplines. Business education covers a range of fields in itself, but there is enough common focus (and common business core curricula in most institutions) to provide some degree of similarity in examining assessment methods in the fields constituting the study of business.

On behalf of AIR, I would like to thank AACSB International for its co-sponsorship of the volumes on assessment in the business disciplines. In addition, a huge thanks goes to Dr. Kathryn Martell of Montclair State University for her Herculean efforts in recruiting chapter authors in a wide range of institutions and business fields and editing the volumes. Only those who have attempted such a feat can truly appreciate the effort involved.

We hope that you enjoy this and the volumes that follow and that you will glean a great deal of useful information about assessment and then apply that knowledge in your programs and classes.

Dr. John A. Muffo
Ohio Board of Regents

Before we can "walk the assessment walk," we must know how to "talk the assessment talk." This chapter is intended as a primer on assessment terminology, the assessment process, and AACSB expectations regarding assessment of student learning. Recent survey results provide the basis for a discussion of current assessment practice in business schools, and how they match with the recently ratified AACSB accreditation standards.

CHAPTER 1
ASSESSMENT IN BUSINESS SCHOOLS: WHAT IT IS, WHERE WE ARE, AND WHERE WE NEED TO GO NOW

Kathryn Martell
Montclair State University
Thomas G. Calderon
University of Akron

Background and Rationale for Assessment

Assessment of learning is currently one of the most frequently discussed topics at colleges and universities in North America and several countries in other regions of the world. What can explain the sudden popularity of this topic? And is it excitement—or anxiety—that we hear underlying the chatter about assessment?

First, we should acknowledge that assessment, both as a topic in education circles and as matter of public policy, has been around for more than two decades. During the 1980s, there was widespread debate about the quality of learning at higher education institutions. Much of the discussion in the early to mid-1980s focused on the failure of higher education institutions as centers of learning. In response, state legislators and the federal government took several steps to make higher education more responsible and accountable for student learning and institutional effectiveness. In 1987, for example, the U.S. Department of Education mandated assessment of institutional effectiveness as a criterion for recognition of regional accreditation agencies. Similarly, many state governments developed similar mandates directed at state supported higher learning institutions as a condition for funding.

In response, the late 1980s and early 1990s saw new emphasis on outcomes in assessing effectiveness of higher learning institutions. Regional accreditation agencies mandated assessment of institutional effectiveness, and in 1991 assessment became an integral part of AACSB's mission-focused accreditation standards. To help universities comply with these new requirements from the AACSB and others now, offices of Institutional Assessment and/or Research sprung up on campuses everywhere to document performance on various performance measures, including admission

standards, retention and graduation statistics, graduates' employment, deployment of financial resources, curriculum design and management, routine strategic planning processes, faculty intellectual contributions and, in some cases, measures of the university's reputation in its community. The expectation was that universities that charted their direction through a mission statement, and developed appropriate processes to achieve its mission, would produce quality outputs—the most important of which would be its graduates. Stakeholders, including legislators, demanded this reassurance.

What has changed in recent years—and propelled assessment back to the top of the higher education agenda—is the form this reassurance must take. Growing public dissatisfaction with the quality of college graduates has led accrediting bodies to call into question the efficacy of the institutional effectiveness measures that were commonly used in the 1990s. The question re-emerged, this time with even more urgency: *How do we know what (or if!) students are learning?* Monitoring the curriculum, faculty qualifications, admission standards, planning processes, and resource deployment clearly wasn't sufficient to ensure that adequate learning was taking place. This led to a demand for "hard evidence" (often in the form of *direct* output measures) that students were developing the skills and knowledge base that the curriculum purported to teach. This new accreditation imperative—which very few universities and colleges had processes in place to meet—has led to a surge of interest in how to collect evidence that university students are, in fact, learning what they are taught.

What It Is: An Assessment Primer

Throughout these volumes, our discussion of assessment focuses on assuring that students are achieving the stated learning goals and objectives of their educational programs. As defined by Tom Angelo (1995, p. 7), assessment is "an ongoing process aimed at understanding and improving student learning." Assessment requires administrators and faculty to (a) identify learning goals and objectives for programs and courses; (b) set meaningful expectations and make them public; (c) systematically gather, analyze, and interpret evidence to determine how well performance matches those expectations and standards; and (d) use the resulting information to document, explain, and improve performance.

While there are many definitions of assessment in the literature, our definition, which is adapted from Angelo's work, emphasizes assessment *as a continuous, systematic process, the goal of which is to improve the quality of student learning.* The process calls for planning, discussion, consensus building, and reflection in addition to measuring, analyzing, and improving performance. Assessment requires a considerable amount of faculty involvement, particularly at the beginning when learning goals are established, and at the end where the data are used to improve the curriculum. These

tasks may also require participation, in varying degrees depending on a program's mission, from such external stakeholders as employers and key alumni.

Reflection is a necessary part of assessment—it is an essential element of each phase of the assessment process. It is through reflection that faculty, administrators, and key external stakeholders consider contextual and background information from multiple sources, utilize specific data and information about their students learning experiences, paint a holistic picture of what's going on, ask difficult questions, and find creative solutions that help to improve learning and close the loop.

Key Assessment Language

Many terms that are not often used in the day-to-day faculty conversations are currently part of the language of assessment. Assessment experts frequently talk about those terms without recognizing that the many faculty may not fully appreciate what they mean (Leskes, 2002). For the most part, assessment practitioners are developing their own language that consists of many seemingly ordinary words that few understand. While these words have specific meaning to assessment experts, they do not always mean the same thing to others. It is not uncommon to find some language differences even among assessment experts. To help the reader navigate through the assessment waters, in this section we identify some of the terminology that is often used in assessment discourses and discuss them briefly.

Learning Goals, Objectives and Traits

Learning goals are the product of faculty reflection on the skills, attitudes, and knowledge that they expect students to learn as a result of matriculating through their institution's programs. They are the roadmap for the curriculum, and are the foundation on which the assessment program is built. However, learning goals are broad and not sufficiently specific and observable to be measured. They are, to quote Trudy Banta, "ineffable." Objectives, on the other hand, are clear statements about outcomes that faculty expect from students. **Objectives** identify specific, observable behaviors and actions related to a goal that faculty will use to describe, monitor, and assess student achievement. Thus, objectives are used as indicators of goals.

To illustrate the difference between goals and objectives, assume that among the goals of a program whose mission is to prepare competent and responsible business professionals is the statement, "Our graduates will be interdisciplinary thinkers." This statement is broad, and not-sufficiently specific and observable to be measured. For assessment purposes, the faculty must define a specific, observable behavior that offers insight into what the statement actually means, and the type of evidence they could use to monitor achievement

and progress. Thus, as a learning objective, the faculty may come up with a statement such as, "When confronted with a case from one discipline, students will appropriately apply methods from another discipline." This statement is specific and the outcome is observable. It tells what students will do, and specifics the expectation that the business curriculum will provide students with opportunities to acquire analytical skills in various disciplines and know when to use them. Another example of a learning goal, "Our students will be ethical," may be translated into learning objectives related to ethical knowledge, recognition/awareness, decision making, or behavior. It is in the process of translating a goal into an objective that faculty must determine what they *really* mean by the goals they have established for their students. Some learning goals—for example, critical thinking, ethics, creativity and leadership—are more difficult to translate into objectives than some others (e.g., communication, knowledge about business, technical proficiency). The section of this volume, *Assessing Business Knowledge and Skills,* provides insight on how to translate learning goals into observable objectives.

Although objectives are designed to be observable, faculty must still develop a measurement process that allows them to measure, monitor, and reflect on the outcome. Measurement in assessment refers to a systematic process of deriving quantitative or qualitative facts about an objective. The measurement process must include identification of very specific characteristics or **traits** (Walvoord & Anderson, 1998) that define the performance-related facts about a learning objective that faculty will use for assessment. Breaking objectives into traits is an essential part of the measurement process, as is establishing standards of performance for each trait. For example, faculty at a Midwest university developed the following traits to assess the objective, "*Our graduates will demonstrate effective written communication skills as applied to a topic in business:*"

(In his or her paper, the student:)

1. Provides a clear introduction and background.
2. Demonstrates knowledge and understanding of the discipline-specific concepts and issues.
3. Uses internally consistent arguments.
4. Organizes arguments in a logical manner.
5. Draws conclusions that are consistent with arguments and analysis.
6. Uses an acceptable writing style and grammar.
7. Demonstrates effective literature search skills.
8. Documents sources used effectively.

By themselves, these traits are not very meaningful. However, when used as the basis for developing performance standards and expectations for student learning, they become powerful elements in the assessment process.

Faculty can use the traits to direct student learning and to develop comprehensive rubrics for evaluating, monitoring, and reflecting on student achievement. A **rubric** is a focused, documented set of guidelines, usually in matrix form, that faculty can use to evaluate student work and provide feedback. As an example, Table 1 provides part of the rubric for assessing written communication skills corresponding with the traits listed above. Additional rubrics are also provided elsewhere in these volumes.

Direct Assessment and Indirect Assessment

There are two basic approaches to measuring learning in assessment: direct and indirect. Despite the apparent simplicity of these measurement concepts, there is a great deal of confusion about what they mean.

With the **indirect** approach, students, alumni and/or employers are asked to provide their opinions regarding the learning that takes place in the school's programs. The popular techniques for collecting these impressions include surveys (students, alumni, employers), focus groups, and exit interviews. In contrast, the **direct** approach bases assessment on students' *demonstrations* of their knowledge or skills. Assessment of oral communication, for example, might require that students demonstrate their skills in a speech or presentation, while assessment of interdisciplinary thinking may require a case analysis. There are many different direct assessment techniques—some take place in the classroom (course-embedded assessment), while others are part of the program's graduation requirement (demonstration assessment).

Course-embedded assessment is efficient and can produce very effective insight into student learning and the achievement of specific objectives. The distinction between assessment of a student within a course and course embedded program assessment is an important one. Follow-up to our survey of assessment practices in business schools revealed that many respondents *thought* their schools were using course embedded program assessment when, in fact, they were not. In order for a course product (e.g., student exams, paper assignments, speeches, case analyses) to be a legitimate program assessment measure, the evaluation criteria and the performance standards must be determined by a group of faculty (not just the teaching faculty member). The course product must be evaluated using those criteria and standards (using a rubric, for example) and the results must be shared with the faculty for the purposes of improving the curriculum. Only then is a course assignment transformed into a course-embedded program assessment.

While the direct approach uses actual work done by students as evidence of learning, there is still debate about the nature of that work. Many assessment practitioners believe the work examined should be based on intellectually challenging tasks with "real world" significance. Thus, they distinguish between **authentic** and conventional assessment by defining the

5

Table 1
Example of a Rubric for Assessing Written Communication

Traits	Performance Levels		
	3 (Very Good)	2 (Satisfactory)	1 (Unsatisfactory)
1. Provides clear introduction and background	Provides a coherent introduction and addresses key background issues effectively.	Provides an introduction, presents some but not all of the key background information.	No or unclear introduction, and absence of background information.
2. Demonstrates knowledge and understanding of the discipline-related concepts and issues	Appropriately addresses all key discipline-specific concepts and issues and most of the minor ones.	Appropriately addresses most of the key discipline-specific concepts and issues but omits or inappropriately identifies a few of the minor ones.	Discipline-specific concepts and issues are not identified or they are identified inappropriately.
3. Uses internally consistent arguments	Arguments presented in the paper are consistent and the different parts of the paper are well integrated and consistent.	Most arguments are consistent, and most parts of the paper are integrated and consistent.	Arguments are inconsistent, different parts of the paper do not relate well to each other and are inconsistent.
4. Organizes arguments in a logical manner	All arguments are well organized, flow logically, and are easy to follow.	Most but not all of the arguments are well organized, flow logically, and are easy to follow.	Arguments are poorly organized, do not flow well, and are hard to follow.
5. Draws conclusions that are consistent with arguments and analysis	Very strong connection between analysis and conclusions; there are no inconsistencies.	Reasonable connection between analysis and conclusions; there may be minor inconsistencies.	No or very weak connection between analysis and conclusions.
6. Uses acceptable style and grammar	No or very minimal spelling, grammar, sentence structure, and paragraphing errors.	Few spelling errors, generally appropriate grammar, sentence structure, and paragraphing.	Numerous spelling errors; inappropriate grammar, sentence structure, and paragraphing used throughout the document.

Table 1 (continued)
Example of a Rubric for Assessing Written Communication

Traits	Performance Levels		
	3 (Very Good)	2 (Satisfactory)	1 (Unsatisfactory)
7. Demonstrates effective literature search skills	Uses multiple research sources and makes effective use of current and reliable information; sources are authoritative or come from the professional literature; includes current relevant sources.	Uses more than one research source but demonstrates no attempt to incorporate current information; most sources are authoritative or come from professional literature; sources are generally current and relevant.	No literature or use of inappropriate web-based sources; sources are not authoritative or do not come from professional literature; sources are not current and/or relevant.
8. Documents sources effectively	Includes appropriate citations within the document and lists references that includes full bibliographic information for each citation in the document; citations and references are consistent throughout the document.	Generally includes appropriate citations within the document and lists references at the end of the document; references are generally usable but may not contain the full bibliographic information; citations and references are generally consistent throughout the document.	No or inconsistent references; evidence of plagiarism.

former as assessment that is based on "worthy intellectual tasks" (Wiggins, 1990), in contrast to traditional tests that emphasize efficiency and contrived, dichotomous (right/wrong) responses. Authentic assessment emphasizes higher level learning skills and requires students to demonstrate performance on "real world" tasks, which are usually ambiguous and necessitate integration of ideas and concepts from multiple sources. Wiggins (1990) argues that authentic assessment "improves teaching and learning [because] students have greater clarity about their obligations (and are asked to master more engaging tasks), and teachers can come to believe that assessment results are both meaningful and useful for improving instruction." Practitioners and

researchers often use the term performance assessment or **performance-based assessment** to describe this type of testing (Muraki, Hombo, and Lee, 2000). Although authentic assessment is valued by many assessment experts, this approach is not required to meet AACSB standards.

Closing the Loop

The overarching goal of assessment is to improve the quality of student learning. Since improving student learning is dynamic and continuous, the literature often refers to the assessment process as a loop that moves from identification of program goals and objectives, to measurement and evaluation, to reporting and dialogue, to identification of program improvement opportunities, and ultimately to reflection and change (see Figure 1). Achieving the benefits of all items in the loop, including reflection and change, is referred to as closing the loop. **Closing the loop** implies that the data collected, analyzed, and reported as part of the assessment process are actually used to improve student learning. It is challenging and requires dedication and time. For most business schools, closing the loop will be a multi-year endeavor. It also will be, for many, the most challenging part of the assessment process. Thus, it is not surprising that as many as 42 percent of business school deans participating in our survey report that they have made no significant progress in closing the loop. About 30% express concern about their business school's ability to close the loop in a manner that might satisfy AACSB's assurance of learning standards.[1] While challenging, closing the loop is the most important part of the assessment process and will bear special scrutiny by accreditors. After all, the primary purpose of designing an assessment process is to *use the data to improve student learning.* Collecting data without acting upon it is a waste of resources, and will not advance the school's accreditation case.

Value-Added Assessment versus Performance-Based Assessment

Given that the goal of assessment is to improve student learning, there is a debate about whether assessment should focus on (a) demonstrating that students are performing at the levels expected by a program's faculty (i.e., *performance-based*), or (b) demonstrating that learning improves as students move through different points in a program (i.e., *value added*). **Value-added assessment** seeks to demonstrate that students learn because of the instruction they receive. Value-added assessment normally uses a pre/post design, with the difference between the post and the pre-test scores providing an indication of value-added. Curriculum designed around the principle that students must demonstrate growth in higher level learning between the start and end of the core business curriculum can also serve this purpose. (See, for example, the competency growth plans used by King's College). As noted below, the AACSB does not require value-added assessment, which is generally more sophisticated than performance-based.

8

A Process Model for Assessment[2]

As described in the preceding section, the goal of assessment is to improve student learning. Business schools need to follow a coherent process to achieve this goal. The AACSB's Assessment Resource Center (AACSB, 2004b) states that a well-designed assessment process will include the following five activities:

1. Define learning goals and objectives.
2. Align curriculum with goals.
3. Identify instruments and measures.
4. Collect, analyze, and disseminate assessment data.
5. Use assessment data for continuous improvement.

These activities are depicted in Figure 1 as a process model. The model shows several enablers of effective assessment, such as faculty and administration support; shared values, particularly in relation to the program goals and expectations for student learning; resources to fund and manage assessment (e.g., money, people, and office space); and information technology for storing, processing, reporting, and managing assessment data and formation. Figure 1 also shows that once goals and objectives are defined, the entire process will evolve around curriculum alignment and the assessment methods selected. Goals and objectives drive both the alignment process and the assessment methods that the school chooses. Measurement and analysis, reporting and discussion, use of assessment data to identify improvement opportunities, and reflection and change are all dependent on

Figure 1
The Assessment Process

9
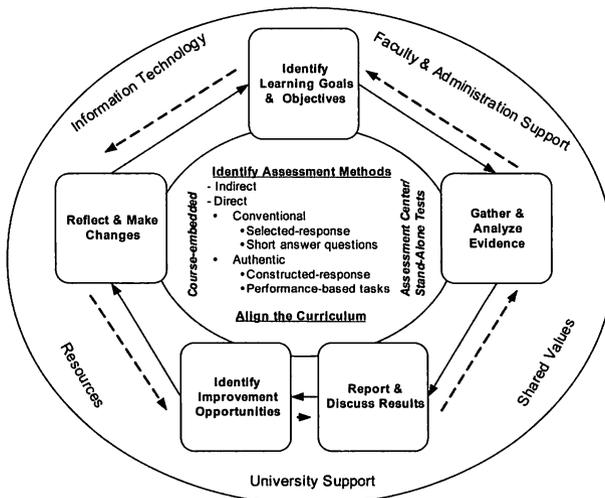
the type of assessment methods selected, and the nature of the curriculum alignment processes. The remainder of this section offers a brief discussion of each major activity in the process.

Defining Learning Goals and Objectives

The necessary first step in developing an assessment program is to define learning goals and objectives. The process needs to start with the goals, not the methods! Learning goals are the general educational aims of the program—the broad outcomes that are expected. Thus, they should flow from the program's mission.

The AACSB requires four to 10 learning goals for each degree program in the business school. Active faculty involvement in defining learning goals is critical and expected. Other constituencies who can meaningfully contribute to the discussion include representatives from the business community, alumni, students, and faculty from outside the business school.

Useful questions to pose to begin the discussion on learning goals include:

- What do we value?
- How would we identify a successful graduate?
- What are the skills, knowledge, experiences, and values that are at the center of our program's curriculum?
- What should a graduate from this program be able to know and do?

Defining learning goals is a thought-provoking, time-consuming, sometimes acrimonious exercise that needs careful attention. These goals will drive not only assessment, but the whole curriculum management process When asked to indicate their top five learning objectives for their undergraduate programs, our survey of business school deans identified: effective communicators (67%), ethics (59%), critical thinkers (42%), effective decision makers (38%), problem solvers (36%), and able to integrate across disciplines (32%). Remember, learning goals should include both knowledge and skills.

Align Curriculum with Goals

After goals are defined and translated into objectives, the next issue to address is curriculum alignment: Where in the curriculum will this learning take place? For each objective, faculty should identify where in the curriculum students will have the opportunity to develop that competency. While the bulk of these educational experiences may take place within the classroom, others may not. In aligning the curriculum with the program's learning goals, both business school and general education courses should be considered, in addition to relevant out-of-class experiences (e.g., required internships). One way to conceptualize this is to complete a grid organized by learning goals. It may be useful to list required courses and other educational experiences along the rows of a grid and list learning goals across the

columns. For each required course, cells can be filled in with learning experiences that build students' competencies in the goal areas.

Not all courses will address all goal areas, of course, but it seems reasonable to assume that all required courses should address at least one key learning objective. The process of developing the curriculum alignment grid may uncover problems relating to inadequate or redundant coverage of specific learning goals—a conclusion that should foster reflection and change even before the assessment process is completed. Table 2 shows an excerpt of a curriculum alignment grid from Montana State University. The excerpt, which lists all courses in the school's business core curriculum but only three of their 10 goals, is interesting because it does not merely employ an "x" to indicate coverage of a goal (as many schools do); instead it offers a brief description of what is covered in the class in relation to the covered goal.

Identify Instruments and Measures

Once the school has aligned curriculum with goals, the next step in the assessment process is to collect evidence to show that students are achieving the program's goals and objectives. There are many different methods that can be utilized to collect this evidence, including course embedded (cases, class exercises, embedded questions on exams[3]) and comprehensive demonstration-based methods (assessment centers, student portfolios, exit exams[4]). The school must also resolve such issues as how, where, and who will administer the assessment.

Appropriate measures and methods must be chosen for each learning objective. Some methods—for example, an assessment center exercise or a senior capstone project—may yield data to evaluate multiple learning objectives. The Senior Assessment panel used at the Stillman School of Business at Seton Hall, for example, generates data that can be used to assess critical thinking, change management, technological competency, and communication skills. Other learning objectives may be relatively difficult to assess—for example, leadership abilities or ethical decision making— and may require multiple activities and measures for a single objective. In most cases, however, a carefully planned activity used with a set of well conceived rubrics (see Table 1) should provide the required evidence of student learning. Schools can aim, over time, for multiple measures.

To meet AACSB requirements, assessment programs need to generate data that provide a basis to evaluate **all** of the program's learning objectives. While a carefully designed capstone project or an assessment center may be capable of this, most schools will find that they need to use multiple methods to generate the breadth of data that is required.

Collect, Analyze and Disseminate Assessment Data

Collecting, analyzing, and disseminating data are shown in Figure 1 as

Table 2
Excerpt from Montana State University's Core Business Curriculum Alignment Grid

Course	Written Communication	Oral Communication	Critical Thinking
101: Freshman Seminar	Memo format, position memo, business plan.	Major focus! Several short presentations, oral midterm, formal presentation of plan.	Problem solving within project context: personal plan, business plan.
201: Managerial Communication	Letters, memos, résumés, quizzes, and formal reports.	Impromptu, extemporaneous, team reports, team discussions, class discussions.	Problem solving, audience analysis, teamwork.
221: Principles of Accounting I	Some assigned problems require a written narrative.	Encourage student response to questions.	Application of accounting principles to business transactions.
222: Principles of Accounting II	Short article reports pertaining to managerial accounting, written narrative problems in homework.	Student presentation of homework problems.	Problem solving applications to current events.
301: Management & Organizations	Case analysis, essay exams, term paper, report, memo. Expectation of few errors.	Oral presentations and class/team discussions.	Introduction to problem solving and decision making models. Analysis and synthesis required in case studies, essay exams, and presentations. Emphasis on rationality and the scientific method.
311: Information Systems	Three group papers, numerous short Email messages, and essay exams.	Small group discussions.	Analyze & evaluate information used in information systems to support decisions.
331: Operations	Seven short group project summaries.	Group interaction through two production simulations and seven group critical thinking projects.	Seven group projects emphasizing problem structuring, information sorting, analysis and summation, and two group production simulations.
341: Marketing	Essay portions of exams and written papers.	Regular article and case discussions.	Model building, case analysis, and synthesis.
351: Finance	Essay portions of exams, 1 group paper, 3 individual papers.	Participative presentations of cases and problem solutions; discussion and debates.	Case analysis and synthesis, problem solving, position debates.
361: Introduction to Law	Essay portion of examinations.	Class discussions and presentations.	Case analysis, reasoning by analogy, synthesis of case material, problem solving.
474: Senior Seminar	Formal professional reports, team authored reports.	Team projects, extensive class discussions, professional level presentation skills.	Focus of course. SWOT analysis, macroeconomic analysis, synthesis of functions, creation of strategy, integration of ethics.

Source: http://www.aacsb.edu/resource_centers/assessment/practices-MSUCurriculumMatrix.pdf.
Last accessed December 28, 2004.

two separate activities—(a) gather and analyze evidence, and (b) report and discuss results. Evidence collection and analysis must be systematic and will reflect the nature of the objective. Thus, for example, evidence of written communication skills is likely to employ a rubric that includes specific, communication-related traits that are valued by the faculty. Similarly, measurement of negotiation skills will be based on traits related to that skill, which are valued by the faculty. Clearly, measurement in relation to performance-based assessment is not an exact science, and there could be instances of scoring variability among different raters (Muraki, Hombo, & Lee, 2000). Schools can overcome this challenge by creating clear rubrics and training faculty to be more effective in recognizing evidence and using rubrics.

In order for assessment to contribute to the curriculum management process, data must be shared with the faculty. Some schools, such as Seton Hall, Rowan, and Missouri Southern State University, prepare an annual report for faculty analyzing assessment data, while others (e.g., King's College) set aside an Assessment Day for faculty to make presentations on assessment results. These reports or presentations then become the basis for evaluating whether the program's curriculum is producing acceptable learning outcomes. Questions discussed at such forums include: What do assessment data suggest? How has the program fared? What is there to celebrate or to be concerned about? How should the assessment data be used to improve the program?

Use Assessment Data for Continuous Improvement

Arguably, the most critical step in the assessment process is the final one—using assessment data to improve school programs. If a significant number of students have failed to demonstrate competence on a key learning goal, a response must be crafted. There are a number of curriculum changes that can be proposed to improve the learning experience, including new course design, revision of the content and/or methodology of existing courses (including courses outside of the business school), course collaboration, and enhancement of out-of-classroom experiences such as tutorials, skill-building workshops, internships, and service learning. Fifty-eight percent of deans we surveyed have made significant curriculum changes because of the (primarily indirect) assessment process.

A second improvement strategy may focus on faculty development, including implementation of teaching workshops on topics that are related to the learning goal (e.g., writing across the curriculum, activity-based learning). Some schools such as Rockhurst University have grant programs that encourage faculty involvement in assessment or curriculum development of targeted skills or content areas.

A final issue to consider after reviewing assessment results is the assessment process itself. Are the learning goals appropriate? Do objectives

need modification? Are the methods and measures generating adequate data for evaluating student competency in each area? It is unlikely that a school's first attempt at developing an assessment process will be its last. Schools that are most effective at assessment use assessment data not only to refine their curriculum, but to refine the process as well. The Rowan University case provides an excellent example of how a school reevaluated and subsequently changed its assessment processes.

Some schools prepare a comprehensive assessment plan to guide them through each major activity in the process. At the most basic level, such plans are prepared in the form of a grid that includes:

- Program learning objectives;
- Assessment methods to be used in collecting evidence about student learning for each objective;
- Identification of the measurement metrics and expectations for each learning goal;
- A time line for implementation; and
- Identification of persons responsible for coordinating data collection.

As the sample form below indicates, these reports do not need to be extensive. A 2-3 page report developed each year for each program summarizing findings and identifying next steps in improving the curriculum and assessment process, followed up with a short progress report on action items the following year, should be more than adequate to keep the assessment process moving forward. The Neely School's example (Vol. 1, No. 2) provides one form of a concise annual assessment report. Another example is provided in Table 3.

Comprehensive assessment plans can be powerful tools for effective management and communication of assessment activities. Just thirty-six percent of business schools that responded to our survey have a formal written assessment plan—we expect this number will approach 100% as the new standards become more widely implemented. In addition, our survey indicates that schools that have a formal plan are more likely than those without to report success in using assessment to improve the curriculum.

What Does the AACSB Require?

The new *AACSB Procedures and Standards for Business Accreditation*—passed with membership vote in April 2003—replaced the 1991 accreditation standards. The 2003 standards are divided into three categories—Participants, Strategic Planning, and Assurance of Learning (AOL)—and it was widely perceived that the AOL standards were a sizeable departure from what was previously required. In some ways, this is true—the

14

Table 3
Example of an Assessment Plan (Excerpt)

Outcomes	Assessment Methodology				Summary of Results	Future Actions
	Measurement Method/Metric	Expectations	Procedures			
1. Students will demonstrate knowledge of core business fundamentals.	**Method:** In-house designed multiple-choice test that covers cognitive outcomes in the core curriculum. **Metric:** Number of correct responses.	Satisfactory total raw score (=60%). Satisfactory performance on each core curriculum learning goal (=60%).	**Description:** Computerized multiple-choice test that covers the domain of the core business curriculum. **Sampling:** Each student registered in the program's senior level capstone course. **When:** Towards the end of each semester **Where:** Senior level capstone course **Incentive:** Students receive 5% of their grade based on performance.	Results are analyzed by item, by each learning objective, and by cognitive learning level assessed based on Bloom. We use tables and graphs. Students are performing well below expectations. This is true of both the overall score and the specific learning goals for the core curriculum.	1. Review and discuss evidence at scheduled assessment review meeting 2. Have faculty who contribute directly to specific learning objectives lead the discussion on those objectives. 3. Identify opportunities for change. 4. Take necessary action to begin implementing at least one change.
2. Students will demonstrate ability to apply core business fundamentals.	**Method:** Embedded comprehensive case/project in senior level capstone course. **Metric:** Performance levels on specific traits in rubric approved by the program's faculty.	Satisfactory rating on each trait listed in rubric.	**Description:** A case/project selected by the faculty that covers the domain of the core curriculum reasonably well. **Sampling:** Each student registered in the program's senior level capstone course. **When:** Towards the end of each semester. **Where:** Senior level capstone course. **Incentive:** Students receive a pre-determined portion of their grade based on this assessment.	This assessment item is still at the planning stage. No results are currently available. We expect to summarize student performance on each trait and share results with faculty graphically and in tabular form.	1. Determine specific criteria for case/project selection. 2. Define traits and performance levels, and complete the rubric by mid-semester. 3. Train faculty to use rubrics and recognize evidence. 4. Agree on proportion of course grade to be allotted to the case/project, and case/project administration logistics. 5. Administer case for the first time next semester.

required documentation has changed significantly (discussed more below) and the topic has received more emphasis than previously when it was folded under "curriculum evaluation." In the past, learning assessment issues were not among the largest obstacles schools faced in achieving and maintaining their accreditation. This will not be the case in the foreseeable future, however. Even as we enter the first year (2005) that the new standards are being fully implemented, a number of schools have already been counseled by their advisors, or told by their peer review teams, that their assessment processes must be improved or they will not receive or maintain AACSB accreditation. So it's a serious topic that has become a priority for business schools, as evidenced by the attendance at AACSB Assessment conferences that have been "sold out" since their inception in Fall 2002.[5]

Before discussing how AACSB's expectations for assessment of

student learning *have* changed, and providing a progress report on how schools are adjusting to the new requirements, let's pause to consider what *hasn't* changed. The AOL standards are consistent with three key aspects of the AACSB philosophy that have been prominent for more than a decade: Standards that emphasize continuous improvement are mission driven, and (related to that) give schools flexibility in how they meet the standards (there is no prescribed "best way").[6] Furthermore, the AACSB's requirements for assessment data are fully consistent with those of other regional accrediting bodies (see chapter 2 of this volume), other professional accrediting bodies (e.g., NASBITE for Education and ABET for engineering), and some state legislatures (e.g., South Carolina, Florida, Kentucky, Michigan, and California). In fact, a number of other accrediting bodies now require more from that their members with regard to documentation of student learning than does the AACSB.[7]

So what *does* the AACSB require? It requires that schools articulate key learning goals for each of their degree programs, assess their students' learning related to those learning goals in a systematic, direct way, and use assessment data to improve student learning. It does *not* require specific levels of performance (e.g., students must score above the 60th percentile on a specified test), that schools compare their students' performance with students in other schools (i.e., benchmark), and does *not* prescribe what percent of students must meet its performance standards. Since faculty (and not the AACSB) choose their school's learning goals and develop their own expectations and performance standards, they are empowered to ask the really important questions—*what do we REALLY want to know about our students' learning?*—without fear of retribution. The goal of the AOL standards is continual improvement; thus, a school's faculty can receive "bad news" about their students' learning (e.g., most of our students cannot write a professional memo!) and not be penalized as long as the school energetically moves to improve performance in that area.

With regard to improved student learning, the focus is on program-level assessment, not individual-level student assessment (like we do in classrooms). Assessments scheduled as students near graduation—which is a prudent time to evaluate the acquisition and retention of skills and knowledge gained during their program—will often not yield information on deficiencies that will be addressed for that group of (graduating) students. The assessment results will, however, inform the school about areas that need strengthening in the curriculum to improve *future* students' learning.[8]

Thus, the goal of the AOL standards is—as it should be!—improved student learning, and that is what schools will be held accountable for at the time of review. The purpose is not to compare schools with each other, or set up a uniform standard of performance that all schools (or students) must meet. Schools have different missions, different stakeholders, and different student bodies, and cannot be meaningfully compared. AACSB has

recognized and celebrated this diversity for years in its "mission-based" standards for accreditation. The 2003 standards are fully consistent with this philosophy.

As we noted above, the starting point of the assessment process is clearly articulated learning goals—what skills and knowledge will our students have as a result of going through our program?—and these are to be (a) mission based; (b) developed for each degree program; and (c) be determined by the faculty. The AACSB does not require goals at the major or concentration level; however, some regional accreditors do (see chapter 2).[9] Each degree program must establish 4 to 10 learning goals and must include goals related to both knowledge and skills.[10] What is a program?

> The school must specify learning goals for each separate degree program. Generally, such goals are anticipated for each degree, not for separate majors or concentrations within a degree. For example, a school may offer a Bachelor of Science in Business Administration (BSBA) degree with defined majors in finance, marketing, human resource management, operations management, and general management. A set of learning goals for the BSBA degree must be provided and assessed; goals for each major (while they may be developed for the school's use) would *not* be required for accreditation review purposes. However, if the school also offers degrees at the undergraduate level called Bachelor of Science in Management Information Systems (BSMIS) and Bachelor of Arts in International Management (BAIM), each of those degrees would require a specification of its learning goals.

> The only exception to this is if the school is seeking separate AACSB accreditation for accounting. Even if accounting is not offered as a separate degree program, the major must have its own assessment system for accounting accreditation purposes. (AACSB Assessment Resource Center, Frequently Asked Questions).

Different degree programs may include common goals (e.g., both the BSBA and MBA programs may have "effective communication" as one of its learning goals), but it is expected that if a curriculum is distinctive enough to constitute a separate degree, there should be some key differences in the learning goals as well.

One of the key differences in AACSB assessment requirements in the 2003 standards with the previous standards has to do with methodology. Previously, the most widely practiced assessments of student learning in

business schools and elsewhere in higher education were primarily indirect measures, especially survey data—students, alumni, and businesses—and actuarial data (graduation rates, retention rates, percent of students with jobs at graduation, etc.). As noted above, indirect measures ask students (or those who know them) to give an opinion about their own learning. For example, the EBI survey asks students, "To what extent did the business program enhance your ability to think critically?" These impressions fall short of providing direct evidence of student learning and do not meet Standard 16, which calls for schools to "... *demonstrate* achievement for key, general management-specific and/or appropriate discipline knowledge and skills that its students achieve in its ... degree program(s)." Now indirect measures are expected to play only a minor role in assessment of student learning, primarily as a secondary measure.[11] Beyond the requirement for direct measures, the AACSB does not prescribe or endorse any particular assessment tool or technique as long as it is appropriate for the learning goal. Finally, with regard to methodology, the AACSB does allow for sampling, as long as an appropriate size and representative sample is used. Be advised, though, that assessment must take place at the individual level; team products are not appropriate assessment measures.[12]

As the name of the standards – "Assurance of Student Learning" – suggests, the point of the assessment process is to evaluate the learning produced by the school's programs, and strengthen curriculum to address areas where students' skills and knowledge fall short of expectations. Reviewers will be looking closely to see how schools use the assessment data they have gathered. Gathering data without subsequently using it to improve student learning is not only a waste of resources, but will lead to serious problems when reviewers question why assessment data were not evaluated, disseminated and acted upon. Thus, schools should take care to choose methods that match its learning goals and generates data that can be readily interpreted, and to develop assessment programs that put an emphasis on "closing the loop."

Where Should We Be Headed?
Elements of Effective Practice

In addition to considering different assessment tasks when developing an assessment process, it is prudent to have some sort of picture of what an effective process might look like. Effective assessment is systematic and carefully planned. Calderon, Green, and Harkness (2004), identified several elements of "best assessment practices," including:

1. Assessment activities focus on a coherent portfolio of shared learning outcomes that align with the mission, goals, and objectives of the program;
2. Assessment involves collection of quantitative and qualitative data

through the use of well defined processes that are likely to have continuity.

3. The assessment program involves and engages more than just a single faculty member, and is supported by the faculty and viewed as a shared responsibility.

4. The program documents its efforts in using assessment data and information to motivate course and program improvement.

5. The assessment program has clear leadership, and senior administrators are committed to and support assessment activities.

6. Program improvement is an overarching, unambiguous goal of assessment.

These best practice elements can serve as a useful checklist for gauging a school's progress on assessment.

Levels of Assessment Implementation

The nature of effective assessment practices implies that implementation will take time. Assessment implementation exists at different levels, and progress from one level to another can be slow (Higher Learning Commission, 2002). Through a content analysis of accreditation reports prepared by several institutions of higher learning, the Higher Learning Commission (HLC) of the North Carolina Association of Colleges and Schools identified three levels of assessment implementation, which can be readily adapted to business programs.

The most elementary level can be described as an **awareness** stage in which faculty and administrators recognize the need for assessment and begin the process of identifying mission, goals, specific objectives, and assessment methods that focus on student learning. A school at this level is beginning to create an initial plan and infrastructure to support assessment. At this stage, AACSB and regional accreditation are still viewed as the primary driver of the assessment initiative, and the language of assessment is still relatively unknown. However, largely because of AACSB's Assurance of Learning Standards, the focus of assessment is beginning to shift from traditional measures of institutional effectiveness towards student learning outcomes. Similarly, there is an emerging dialogue about deemphasizing indirect assessment methods and emphasizing more deliberate, direct approaches to gathering evidence about student learning. While a shared understanding of the purpose, advantages, and limitations of assessment is beginning to emerge, there is not widespread faculty or student support for assessment. Students, in particular, know little about assessment, and they do not understand their role in assessment activities. According to our survey, most business schools are at this stage now.

The second level can be characterized as an **initial implementation** stage, in which the first assessment plan and many parts of the required

infrastructure are complete. Processes for collecting, storing, retrieving, and reporting assessment data are being implemented. The transition from indirect to direct methods of assessment of student learning is definitely apparent, and there is growing support among faculty for this approach to assessment. Additionally, the language of assessment is understood by more faculty, and dialogue about the meaning of assessment data is beginning to emerge among some faculty groups. At this stage, there are more funding opportunities and other resources as well as clear lines of responsibility for assessment. However, assessment data are not being used effectively across the school to motivate program improvement. The business school is only beginning to create a culture of assessment. This stage may actually be spread over several years as the school implements various parts of its assessment plan, makes students more aware of their role in the process, and changes faculty perceptions, attitudes and approaches to assessment. The AACSB's "Transition to Assurance of Learning Standards" states that business schools seeking to maintain their accreditation should be at this phase very soon— data collection efforts for at least some learning goals should be under way by 2005 (see "Transition to Implement Assurance of Learning" timetable, page 68 of the Accreditations Standards, revised in January 2004). *In other words, to meet AACSB expectations, this is where business schools should be now.*

At the third level, characterized as a **mature implementation stage**, assessment is a significant priority and is engrained in the school's philosophy and culture. Student learning and continuous improvement are now at the center of the school's culture. Students understand their role in assessment and may even have a representative on the school's assessment committee. Access to assessment data exists seamlessly through the school's intranet or other online resources (see, for example, the chapter on Berry College in Vol. 1, No. 2). Faculty and administrators discuss assessment data regularly, and evidence about student learning is now a consistently important part of the input used in making curriculum and course related decisions. Funding to support assessment is a line item in the school's budget, and the school has an individual with responsibility for directing and coordinating assessment. Fundamentally, closing the loop is now routine and systematic. *The Seton Hall, Kings College, SIUE, Texas Christian, and Rowan cases presented in subsequent chapters of this book are good illustrations of assessment programs developed over time and which now are at the mature implementation stage.*

At many schools, progress through the three implementation levels may not be linear. In fact, results from our survey of business school deans suggest that this might indeed be the case at many schools. Sixty-nine percent of the schools we surveyed have established learning goals for their programs, 34 percent have translated those goals into specific learning objectives, and 36 percent have a formal assessment plan. However, 70 percent

say that they discuss assessment data regularly with their faculty. These statistics hint at a non-linear approach to assessment implementation. Many schools may not develop and pursue the types of formal processes we discuss in this chapter in a sequential and orderly fashion. For example, the University of Houston provides a model of a school that made a large investment in a relatively short period of time (one year) and developed an assessment program that is reaching mature status very quickly. It is also possible that some schools are, in many respects, still at the first implementation level and are struggling with the language, culture, and processes of assessment despite having started on the process some time ago.

Where Are We? The Current Status of Assessment in Business Schools

Our recent survey of business school deans reveals that not only is there a rather sizeable gap between actual and "best" practice, but that many schools are lagging behind the "transition to assurance of learning goals" timetable set by the AACSB. According to the accreditation guidelines, by 2004/05, the AACSB expectation was that schools would have their learning goals fully operationalized, and be experimenting with methods. Although 68% of respondents indicated their schools did have general learning goals, only 31% had operationalized them (i.e., translated them into objectives), and just 17% had established performance standards—all of which are steps associated with the earliest steps in the assessment process. The transition to AOL timeline was established to allow time for faculty to discuss goals, choose methods, gain experience with assessing one or two goals, and apply "lessons learned" to the remainder of its goals. The timeline started at the time the standards were accepted by the membership (April 2003), and it would appear that many schools let a year or more lapse without activity. The complete assessment process incorporating all learning goals, and using assessment data to improve the curriculum, is expected by 2007. Schools that face reviews in 2006 and 2007 need to push quickly to move their assessment processes forward if they have not already done so.

The second problem evident in the survey data is that there is still a focus on indirect assessment measures such as EBI and other surveys. Eighty-one percent (81%) of the schools surveyed are using student surveys, 75% alumni surveys, and 58% use employer surveys. By far, the most popular direct measure reported in our study is the ETS field test – a relatively expensive method that, as discussed in Chapter 6, yields data that some schools have trouble using to improve their curriculum. One of the most common misperceptions we have seen from deans regarding assessment is a misplaced sense of confidence/satisfaction in an assessment process built largely on surveys, interviews, and input from their advisory boards (all indirect measures). While every assessment method used does not have to be original, the AACSB expectation is that "schools (will) choose, create,

and innovate learning measures that fit with the goals of the degree programs (and) pedagogies in use." Our study indicates that few schools have developed their own measures. (Some excellent examples of original measures are provided in case examples in this series.)

The top ranked concern voiced by deans in our survey was finding the resources to conduct assessment, and our data suggest that assessment programs are indeed underfunded. Two-thirds of respondents indicated that some financial resources were devoted to learning assessment in their business schools; however, the amounts were low (75% reported less than $5000), and only 10% have assessment as a line item budget categoryin the business school budget. To put this in perspective, the budget for assessment in most business schools is less than one month's salary for a single junior faculty member. In late 2004, just 26% of respondents report having some release time allocated to assessment (on average two courses per year). On the positive side, 60% report some university support for assessment activities, most notably for faculty training, assessment instruments, and books and materials.

As might be expected in this period of transition to new AOL requirements, many schools seem to be struggling with how to organize the assessment process. Responsibility for assessment has not been clarified in many schools, and currently is spread across multiple committees (11% which are newly formed), department heads, and the dean's office (most popular response). Only 3% of respondents have a staff person assigned to assessment (however, this number appears to be rising). There also is very little experience with the later steps in the assessment process. Record-keeping of assessment data is unsophisticated (75% relying on paper files). Assessment data are disseminated most often as part of regularly scheduled faculty and committee meetings. In time, we expect that more sophisticated and routine processes will emerge.

Given these apparent difficulties in getting started, it was somewhat surprising that many deans in our survey expressed confidence—even enthusiasm—about the assessment task (42% were very confident in their school's ability to design an assessment plan using direct measures, and 45% were very confident of their school's ability to analyze, disseminate and store assessment data). This confidence may be based on success with using (primarily indirect) assessment data in the past. Seventy-nine percent reported their schools used assessment data to make curriculum changes in the past three years, including changes in program requirements (47%), pre-requisites (45%), and course content in the core (51%) and in the major (37%).

As mentioned above, deans' areas of greatest concern in meeting AOL standards are the financial and time resources it will require, reviewers' interpretation of the standards, and that AACSB expectations will change again. Interestingly, the lowest level of concern was expressed about "what

we'll find out," which may also explain the fairly high level of confidence (50%) reflected in their school's ability to "close the loop." Sixty-two percent (62%) of respondents expressed moderate satisfaction with their school's assessment process, 3% were very satisfied, and 35% were dissatisfied. Additional discussion of some obstacles to assessment, and recommendations on how to overcome them, is provided in the final chapter in this series (Vol. 1, No. 2).

Closing Thoughts: Where We Need to Go

Experts often recommend moving slowly with regard to assessment – trying one goal or method at a time in order to gain confidence, experience and faculty buy-in. Unfortunately, this approach is becoming less viable as accreditation deadlines loom ahead. In 2003, the AACSB provided a timeline that allowed schools more than four years to design and implement a full assessment program; however, many schools have let half of that time slip away. Some deadlines by other bodies are even closer. The Florida Learning Compact, for example, which was passed in March 2005, demands full assessment processes to be in place in colleges within their jurisdiction within six months —all the more problematic since half the time allowed for implementation is over the summer months. Needless to say, many Florida colleges and universities are now scrambling to meet this deadline (or risk losing their funding).

Thus, moving slowly has become a luxury that many schools cannot afford. Given the gap between actual and required practice, we strongly recommend that schools in the earliest stages of implementation need to get moving—now! By now, goals should be established, instruments chosen, an assessment plan developed, and data collection begun. These tasks must become a top priority, especially for schools facing accreditation or maintenance visits in 2006 or 2007.

Given these time pressures, it may be tempting to adopt goals that are easy to measure, or methods that are relatively easy to use. This temptation must be resisted. If assessment programs are not built around learning goals that are important to the faculty, and do not use methods that match those goals, it will quickly become a meaningless exercise resulting in cynical faculty, wasted resources, and no basis for improved student learning. As Doug Eder (author of chapter 4) puts it so well, "It takes just as much effort to answer bad questions as it takes to answer good questions." Ask good questions, and honestly seek their answers.

With regard to choosing methods, we advise deans to forget about surveys and other indirect measures when thinking about assessment. As we previously stated, survey data can play an important role in keeping curriculum current, improving student services, etc., but it has very little evidentiary value for assessment of student learning. In our research for this book, we talked with many deans and often found a sense of confidence

(even pride) in an assessment program built around surveys that would not maintain their AACSB accreditation today. Given the time needed to put an assessment process in place—and cultivate the faculty "buy in" needed to make it work—there is no time for false confidence.

Our final recommendation on designing assessment programs is to keep them simple. Assessment experts make the distinction between "scholarly rigor" and "academic rigor." Scholarly research has a set of standards (proven methods, replicated results, scientific sampling) that is not usually appropriate for program assessment of student learning. Not only are these standards often impossible to meet since assessment, as a field of inquiry, is still in its infancy, but demands for rigor can stall—even strangle—progress. An honest effort to investigate students learning through direct measures is what is required—not meeting standards for peer-reviewed research. It is much more important to get started, knowing that there is room to improve as a result of experience, than waiting for the perfect assessment to come along.

Finally, we must try not to lose sight of the purpose of assessment. While accreditation might be the imperative to get us started, it is not the only or even the best reason for devoting resources to assessment. The purpose of assessment is to do our job as educators even better – to monitor our institution's most important "output," our students' learning, so we can facilitate it as best we can. To quote Thomas Angelo, former director of the AAHE Assessment Forum: "Accountability matters, but learning still matters most" (Angelo, 1999). Effective assessment programs give us the opportunity to pursue both.

Endnotes
1   Based on a survey of 179 business school deans (70% from AACSB accredited schools) we conducted in Fall 2004. Additional results from this survey appear later in this chapter.
2   This section is drawn from the overview to the assessment process published on the AACSB's Assessment Resource Center Web site, developed by one of this chapter 's authors.
3   Eastern Kentucky, King's College and Cal State Fullerton provide examples of course embedded assessment.
4   See Merrimack College, Valparasio and Seton Hall for examples of comprehensive, demonstration methods.
5   Over 600 faculty and administrators from more than 200 schools have attended the AACSB Assessment seminars from 2002-2005.
6   "...certain standards or portions of the standards apply differentially, depending on the various missions and objectives of different schools," 1991 Standards for Business Accreditation, page 54
7   For a discussion of regional accreditors' demands for AOL data, see Chapter
8   Program assessment at King's, Seton Hall, Valparaiso, and TCCU evaluate students at the beginning and end of their programs; thus, they are able to

use their programs for student development as well. These schools have "mature" and well funded assessment programs. Most schools with little or no assessment experience should not start with this approach.

9   It is a good idea to consider all demands for assessment at the time that an assessment plan is developed in order to collect and store data that can serve multiple purposes

10   While normally all of the topics and skills listed in Standard 15 and 17 are expected to be included in the business curriculum, it is not necessary (or recommended!) to assess students' competency in each topic and skill area.

11   Surveys still have value in evaluating student satisfaction with student services and faculty interaction (included in the participants standards), and alumni and business surveys are certainly useful in keeping curriculum current. Surveys, though, are just no longer seen as adequate measures to assess student learning.

12   The exception to this prohibition of team products would be if they were used as part of an assessment of teamwork.

References

AACSB. (2004a). *Eligibility procedures and standards for business accreditation*. Tampa, FL: AACSB. Retrieved January 11, 2005, from http://aacsb.edu/accreditation/standards.asp.

AACSB. (2004b). Overview of assessment. Tampa, FL: AACSB. Retrieved December 28, 2004, from http://www.aacsb.edu/resource_centers/assessment/overview-process.asp

Angelo, T.A. (1995). Reassessing assessment. *AAHE Bulletin*, 47 (8), 10-13.

Angelo, T. A. (1999. May) Doing assessment as if student learning matters most. *AAHE Bulletin*. 51 (9), 3-6.

Calderon, T. G., Green, B.P. & Harkness, M. (2004). *Best practices in accounting program assessment.* Sarasota, FL: American Accounting Association.

Higher Learning Commission. (2002). Assessment of student academic achievement: levels of implementation. *Addendum to the Handbook of Accreditation* (2nd ed). (March). Chicago, IL: North Central Association of Colleges and Schools.

Leskes, A. (2002, Winter/Spring). Beyond confusion: An assessment glossary. *Peer Review.* Retrieved January 11, 2005, from http://www.aacu-edu.org/peerreview/pr-sp02/pr-sp02reality.cfm

Muraki, E. Hombo, C. M. & Lee Y. (2000). Equating and linking of performance assessments. *Applied Psychological Measurement, 24*(4), 325–337.

Walvoord, B. E., and Anderson, V. J. (1988). *Effective grading: A tool for learning and assessment.* San Francisco: Jossey-Bass.

Wiggins, G. (1990). The case for authentic assessment. *Practical Assessment Research & Evaluation* 2(2). Retrieved December 26, 2004, from http://PAREonline.net/getvn.asp?v=2&n=2 .

Author(s) Bio

Kathryn Martell, Ph.D., is Associate Dean and Professor of Management at the School of Business at Montclair State University. More than 600 faculty from more 200 universities have attended the seminars that she teaches with Dr. Doug Eder (SIUE) for the AACSB on Assessment of Student Learning. She is also a frequent speaker on assessment at AACSB national and regional conferences, and developed the content for AACSB's online Assessment Resource Center (www.aacsb.edu/ARC). As a teacher, curriculum projects she created using technology to link students in the US with students in other countries received recognition from the AACSB (Midwest Regional Award for Curriculum Innovation), ACE (Special Recognition for Cost Effective Curriculum Programs), AAUA (International Award for Innovation in Higher Education), and the U.S. Department of Education.

Thomas G. Calderon, Ph.D., is Professor of Accountancy and Director of College of Business Administration Quality Assessment at the University of Akron. He has published and presented several articles in the area of assessment, led workshops on the topic, and recently co-authored a monograph titled *Best Practices in Accounting Program Assessment,* published by the Teaching & Curriculum Section of The American Accounting Association. He is the 2004-05 Chair of the American Accounting Association's Teaching & Curriculum Section.

AACSB standards are not the only accreditation mandates facing business schools in the US; regional accreditation standards at the institution level must be satisfied as well. As external calls for accountability have heated up in the last five years, the assessment of student learning has taken on a much more significant part of the process of determining institutional quality through regional accreditation. Assessment of student learning has emerged as a primary means of determining institutional effectiveness. Each of the regional accreditation agencies has a slightly different approach to the assessment of student learning, but each is now requiring that schools provide evidence not only of plans to assess, but also evidence of student learning itself and actions taken as a result of analysis of that evidence. Business schools that take into account both AACSB and regional accreditation mandates can design an assessment program that can satisfy both.

CHAPTER 2
ASSESSMENT IN THE CONTEXT OF ACCREDITATION
Robert T. Mundhenk
American Association for Higher Education

Background on Regional Accreditation

Unlike the higher education systems and institutions in other developed countries around the world, higher education in the United States has a long tradition of independence and self-regulation. Without a central Ministry of Education to define or monitor performance, U.S. institutions are comparatively unfettered by government regulation or oversight, so they are free to provide curricula, pedagogy, and degrees that are consistent with their individually-defined missions. This freedom is tempered somewhat, depending on the degree of dependence on public funding. Thus, for example, institutions dependent on public funding often find strings attached to those funds. State departments of education might demand the right to approve and review programs offered, for example, and their boards may be at least partially political appointees. In addition, those using federal funds for research or student aid find themselves subject to a number of federal reporting requirements. But while institutions may feel burdened by reporting requirements or limited by government intrusions in governance, there have been no attempts to date to dictate curricular content or instructional quality.

One of the reasons that higher education institutions in the United States have such a relatively free hand in managing themselves is that federal and state governments have ceded oversight of institutional quality to regional

and program-specialized accrediting agencies that, using a combination of self-reporting and peer evaluation, have been able to certify the degree to which institutions are successfully meeting their stated missions. No other nation allows institutions of higher education so much autonomy; no other nation allows institutions to assert their quality and fulfillment of mission on the basis not of objective, publicly disclosed data, but of the assertion of a regional or specialized accrediting agency that they are maintaining quality consistent with the agency's standards.

But as higher education has grown increasingly more expensive and its benefits as compared with its costs have become less clear, demands for more concrete information about quality and success—than the assertion that an institution has met minimum accreditation standards—have increased. Constituents of all sorts—students, parents, governments, employers—are seeking greater accountability. It might even be argued that the success of such publications as the annual college guide of *U.S. News and World Report*, which never address issues of academic quality directly, shows the public increasingly interested in knowing what it gets for its dollars. The era of institutional autonomy and minimal disclosure seems at an end.

One of the most significant indicators of this historic change is the gradual introduction, by all regional accrediting agencies over the past decade, of the requirement that institutions assess and document the degree to which students are achieving the programmatic and institutional learning goals that institutions of higher education claim they have achieved on graduation.

Certification of institutional quality has always been the primary function of regional accrediting agencies, just as certifying program quality has been a function of specialized accrediting bodies like AACSB. Traditionally the standards applied were *input measures*—number of faculty with terminal degrees, faculty publication, student/faculty ratios—that said little, except perhaps by inference, about the quality or nature of student learning, which most people outside the academy see as the primary function of higher education. Scenting this change in the wind, regional accrediting agencies, and some specialized ones, are now requiring proof that stated learning goals—*output measures*—have been achieved as a part of the process of certification, or accreditation.

In many respects, this change is a natural development in the evolution of accreditation agencies and processes. A short glance at that evolution may help in understanding the current emphasis on assessment of learning.

Expanding the Focus of Accreditation
The first regional accreditation agencies grew out of loose regional associations founded between 1885 and 1899.[13] and their purpose was, in a very real sense, exclusionary. As El-Khawas argues, "the standards that were used during this embryonic form of accreditation were established mainly to define membership and to distinguish between secondary-level and college-

level study (2001, p.30). In other words, accreditation was aimed neither at improvement nor at meeting standards, but at setting descriptors and quantitative standards that separated one level of education from another.

Roughly twenty years later, accreditation agencies began to emphasize quality within the context of institutional mission. The North Central Association led the way, deciding that "each institution should be judged qualitatively, on a 'total pattern' of activities and in accordance with its own stated purposes. (el-Khawas, 2001, p.31). This new approach introduced two important additions to the accreditation function: qualitative measures and the context of mission. That context guaranteed institutional autonomy by placing standards within the context of institutional mission, a principle that remains the foundation of accreditation; in the May 2004 "Recommendations for the Reauthorization of the Higher Education Act," the American Council on Education and a number of other higher education associations argued for "the accreditors' responsibility for assuring academic quality and student learning in relation to institutional mission." (American Council on Education, 2004, p.17)

The Congressional committee that received this report was not particularly sympathetic to the assertion, partly because accreditation had moved almost fifty years earlier into a more public phase. Whereas accreditation—both regional and specialized—through the late 1940s tended to involve reporting data and receiving accreditation, in the 1950s the concept of periodic self-study, followed in the late 1960s and early 1970s by a move toward site visits by outside peer evaluators, changed the emphasis of the accreditation process from validation of quality based on data reporting to analysis of data to improve quality and remedy problems. At the same time, the introduction of teams of peer evaluators made the process of accreditation much more public, albeit within the confines of academe. By introducing the idea of external review, however, it opened the doors to the concept that institutions are accountable to entities, and answerable to common standards, outside the institution.

Thus regional accreditation practice, as it evolved, became an oddly public and private enterprise at the same time. Kenneth Young's definition of "voluntary accreditation" illustrates this tension. He sees it as ". . . a concept . . . unique to the United States by which institutions of postsecondary education or professional associations form voluntary, non-governmental organizations to encourage and assist institutions in the evaluation and improvement of their educational quality and to publicly acknowledge those institutions, or units within institutions, that meet or exceed commonly agreed to minimum expectations of educational quality (quoted in Harcleroad, 1980, p.12). Note the tension in this description: the accrediting agency's primary function is to *assist* institutions to *improve* and to *publicly* identify those institutions that meet *minimum* standards. The accrediting agency—an external body, represented by external peer evaluators—examines institutions

and identifies those that meet minimum standards. The process has a public orientation and is based on common standards; it thus challenges and, given the importance of accreditation in institutional marketing and eligibility for governmental dollars, overrides the concept of absolute institutional autonomy.

Assessment and Accreditation

Just at the point at which voluntary accreditation, oriented toward institutional improvement and based on internal self-study and external peer review, took firm hold, educational researchers began exploring the uses of data on specific aspects of student learning as a measure of higher education's success. Instead of looking at acceptance or graduation rates as measures of quality, they tried to look at what students were learning— and how well. They began to ask a few basic questions: What do students know as a result of their experience at a particular institution? What can they do with what they know? How do institutions know and demonstrate that students know and do what institutions claim? Though basic, these questions were not simple. They forced institutions to makes a connection between the claims they made in mission statements and the evidence they collected to demonstrate institutional quality.

That confluence of external demands for greater, more relevant information about what students had learned and could do as a result of a college education with increased internal interest in the process and results of learning, made the connection between accreditation and assessment of learning an obvious one. Though data collection strategies were still fairly crude (and, in some cases, remain so of necessity) in the early- to-mid eighties, institutions and individual researchers began to develop plans and systems that could document learning gain. How this information was to be used, both within the accreditation process and within the institution, was somewhat problematic, however.

Institutions, recognizing the importance of data on learning to their ability to document their own institutional effectiveness, generally attempted to treat assessment processes and data as quality improvement efforts. Those that used these data as inferential data on faculty effectiveness found that the contention over faculty evaluation obscured whatever advantages they might have gained for learning about student learning. (It must be added that the opposition to using student learning data for faculty evaluation, except in extreme cases, had a valid point, for there are too many other variables beyond teaching effectiveness that factor into student learning.) And so accrediting agencies took a more diffident approach, one that respected institutional autonomy and mission. They generally required evidence that assessment of student learning was taking place, that the assessment data were being used to improve the institution and to document institutional effectiveness, and that assessment processes were somehow contained within an institutional plan.

For the last ten years at least, assessment of student learning has been given at least nodding recognition in accreditation criteria and institutional action, but criteria were not particularly directive nor institutional responses particularly profound in most regions But, as Frank Brush Murray points out in an article on accreditation in teacher education:

> The case of teacher education accreditation is a microcosm of higher education accreditation writ large. It is very difficult, for example, to find those who think that American higher education, whether accredited or not, is living up to the trust and confidence the public has invested in it... the charge against higher education is that it is not delivering on its promises... Accredited institutions, in particular, are seen as excessively costly and self-serving while failing to meet their obligations and promises .... we can expect to find pressures, similar to those found in teacher education, on the accreditation mechanisms in other professions. They too will be called on to provide solid evidence that their members are fully competent and qualified if they are to extricate themselves from intrusive and misplaced oversight by other bodies (2001, pp.62-63).

External calls for accountability—the "pressure" to which Murray refers—have heated up in the last five years, though, both accreditation agencies and institutions are seeing the assessment of student learning as a much more significant part of the process of determining institutional quality through accreditation.

Each of the regional accreditation agencies has a slightly different approach to the assessment of student learning, but each is now requiring that schools provide evidence not only of plans to assess, but also evidence of student learning itself and actions taken as a result of analysis of that evidence. Assessment of student learning is now a primary means of determining institutional effectiveness. While the approaches to assessment may differ region to region, the core requirement that institutions assess student learning remains constant. In the following brief survey of the policies and standards of each of the regional accrediting agencies, one thing is clear: assessment of student learning is a central element in all accrediting processes across the country.

The Southern Association of Colleges and Schools

The Southern Association of Colleges and Schools (SACS) was one of the earliest regional accrediting agencies to require evidence of gains in student learning as a part of its accreditation processes. Though its standards were (and remain) relatively terse on the issue, two of its members published a

guide to assessment for SACS members called *Assessment in Accreditation.* The premises underlying the guide reinforce the theme that assessment has become an important part of accreditation at least partly because of external pressure:

> There is another reason for the skepticism of public officials about higher education's claims for autonomy. They have observed that the self-regulation of professionals in any field, whether it be health, the law, or higher education, often puts the self-interest of the profession above the public interest, and some oversight by a body that puts the public interest first is necessary. Higher education has accepted the principle of a limited oversight role by the public without giving up its claims and interest in being largely self-regulating (Folger and Harris, 1989, p. 2)

Folger and Harris argue here that institutional effectiveness must be *demonstrated* externally, not merely *ratified* internally—that is, within a closed professional circle. They acknowledge that "development of ongoing assessment of results of institutional programs involves a substantial change in the accreditation process" (1989, p. 9) and that it "augments the traditional practice of evaluating institutions by their conformity to accepted academic conventions of resources and processes." (1989, p.11)

This gloss on the SACS approach might just as easily be applied to most of the other accreditation agencies. It emphasizes the shift from descriptive input data to value-laden output data, the broadened audience for accreditation, and the accompanying, necessary, "substantial change in the accreditation process." But while the processes of accreditation are changing, the articulation and implementation of these changes vary region to region.

Within SACS standards, for example, the emphasis on assessment results and processes is clear:

> 3.3.1 The institution identifies expected outcomes for its educational programs and its administrative and educational support services; assesses whether it achieves these outcomes; and provides evidence of improvement based on analysis of those results (2004).

This is the sole standard in a section called "Institutional Effectiveness," and its solitary placement suggests its importance. Assessment processes are the means by which institutions demonstrate their effectiveness; there are no other alternatives suggested.

A similar terse approach is taken to the undergraduate curriculum:

3.5.1 The institution identifies college-level competencies within the general education core and provides evidence that graduates have attained these competencies (2004).

The critical thing to note here is the shift to competencies, however an institution may choose to define them. Enrollments, persistence, retention, graduation rates, average grade point averages—all of which were once used as indirect indicators of program quality—are not even mentioned in these standards. This is not to suggest that institutions are being discouraged by SACS from using such data, because they are not; rather, it suggests that institutions are now expected to demonstrate their value by proving their effect on their students.

This approach is not unique to SACS. Every other regional accrediting agency uses similar, outcomes-oriented language. They differ largely in the degree to which they define the outcomes to be examined and in the ways in which that examination is incorporated into institutional processes.

The Northeast Association of Schools and Colleges

The Northeast Association's Commission on Institutions of Higher Education (because of the audience of this book, only those commissions dealing with baccalaureate degree institutions are covered) has published draft criteria that are much more extensive and specific than the SACS criteria. Thus, for example, its standard 4.43 defines fairly explicitly what a NEASC assessment program should contain:

The institution implements a systematic and broad-based approach to the assessment of student learning focused on educational improvement through understanding what and how students are learning through their academic program and, as appropriate, through experiences outside the classroom. This approach is based on a clear statement or statements of what students are expected to gain, achieve, demonstrate, or know by the time they complete their academic program. The approach provides useful information to help the institution understand what and how students are learning, improve the experiences provided for students, and assure that the level of student achievement is appropriate for the degree awarded. Institutional support is provided for these activities (2004).

Note that, whereas the SACS standard simply required that a system for assessment of outcomes be in place, the NEASC standard goes further, defining the nature of the outcomes ("what and how students are learning"), requiring specific statements of program outcomes ("a clear statement or statements of what students are expected to gain"), and defining the degree

in terms of student achievement, not credit hours. While the institution retains significant autonomy in terms of program design, pedagogy, and degree definition, its autonomy is clearly no longer absolute.

NEASC's requirements for undergraduate programs go no farther than requiring the specific articulation of assessable outcomes, but its statement on the assessment of general education (Standard 4.18) is much more explicit and directive:

> Graduates successfully completing an undergraduate program demonstrate competence in written and oral communication in English; the ability for scientific and quantitative reasoning, for critical analysis and logical thinking; and the capability for continuing learning, including the skills of information literacy. They also demonstrate knowledge and understanding of scientific, historical, and social phenomena, and a knowledge and appreciation of the aesthetic and ethical dimensions of humankind. In addition, graduates demonstrate an in-depth understanding of an area of knowledge or practice, its principal information resources, and its interrelatedness with other areas (2004).

This is an excellent description of what a college graduate ought to be, know, and do, and it is the kind of statement that often appears in institutional mission statements. What makes it remarkable here is that it is an accreditation standard, one with which every institution in the NEASC region must attempt to comply, and to demonstrate its compliance through "a systematic, broad-based approach to the assessment of student learning."

Commission on Higher Education, Middle States Association
The Middle States Association's approach to assessment eschews specific statements about curriculum elements but is very explicit about the kinds of procedural information an institution must have to comply with two deliberately interrelated standards, Standard 7 (Institutional Assessment) and Standard 14 (Assessment of Student Learning). In its contextual statement about Institutional Effectiveness, the Middle States Commission makes the centrality of the assessment of student learning outcomes assessment absolutely explicit:

> Information obtained through assessment should be used as a basis for assessing the institution's effectiveness in achieving its stated goals. In addition, outcomes assessment should be linked to an institution's ongoing planning and resource allocation processes. Consequently, evidence gathered about students' development and learning outcomes can be used

to make judgments about resource allocation in planning for overall institutional effectiveness and to enhance academic programs (2004, p. 21).

Thus the connection between effectiveness in achieving student learning outcomes and institutional effectiveness has been made explicit; in turn, the role of outcomes assessment in institutional planning and budgeting is clear.

In Standard 14, the link between institutional mission and student learning outcomes is delineated, and specific requirements based on that link are articulated:

The mission of the institution provides focus and direction to its outcomes assessment plan, and the plan should show how the institution translates its mission into learning goals and objectives. In order to carry out meaningful assessment activities, institutions must articulate statements of expected student learning at the institutional, program, and individual course levels, although the level of specificity will be greater at the course level… statements of expected student learning must be available on campus to those planning or implementing assessment activities and to those evaluating the institution… The institution should specify those assessment measures, methods, and analyses that will be used to validate stated expectations for student learning (2004, pp. 50-51).

The use of "must" and "should" throughout this passage makes it relatively clear that Middle States sees the articulation of student outcomes at course, program and institutional levels; the existence of an assessment plan; evidence of the use of assessment information to improve learning; and the use of learning assessment to document institutional effectiveness as necessary parts of an institution's dossier if it is to receive accreditation. In fact, that view is made incontrovertible in the statement that "an accredited institution is characterized by" these four elements; presumably lacking one or more of them would have some negative effect on accreditation.

The Middle States approach is fairly stringent on the articulation and documentation of goals and their achievement, less so perhaps on the specific components of a college degree than NEASC seemed. Yet, in a monograph called "Student Learning Assessment: Options and Resources," the Commission on Higher Education expects of an institution "that its curricula are designed so that students demonstrate college-level proficiency in general education and essential skills, including oral and written communication, scientific and quantitative reasoning, critical analysis and reasoning, technological competence, and information literacy. . . " (2003, p. 1). Thus,

like NEASC, Middle States expects that curricula will be organized in such a way that certain general skills will be developed and ultimately assessed.

The three remaining accrediting agencies have reduced the number of standards applied to their member institutions in order to emphasize key functions. In each, the effectiveness of educational offerings and, by implication, the assessment processes necessary to document it, are essential elements.

The Northwestern Commission on Colleges and Universities

Standard 2.B of NWCCU's accreditation standards contains three major requirements: that its processes for assessing student learning are clearly defined and integrated into other institutional processes, like planning; that expected learning outcomes are published and regularly assessed; and that the results of assessment "lead to the improvement of teaching and learning." These fairly straightforward standards are undergirded by Policy 2.2: Educational Assessment, which indicates early its emphasis on student learning outcomes: "The communities of interest served by the accreditation enterprise have come to appreciate the validity and usefulness of using output evaluations as assessment as well as input measures." The Commission is clear that it does not wish to impose a particular form of outcomes assessment on its members, but it also directly states that:

> The intent of Commission policy is to stress outcomes assessment as an essential part of the ongoing institutional self-study and accreditation processes, to underline the necessity for each institution to formulate a plan which provides for a series of outcome measures that are internally consistent and in accord with its mission and structure, and, finally, to provide some examples of a variety of successful plans for assessing educational outcomes.

This statement is followed by a list of "illustrative and exemplary" outcomes measures that "could yield an efficacious program of outcomes assessment," among them student enrollment and persistence data, mid-program assessments, end-of-program assessments, program review, and

In a Web-based article called "Evaluating Quality and Effectiveness: Regional Accreditation Principles and Practices," which concerns regional accreditation in general, Ronald L. Baker, Associate Executive Director at NWCCU, seems to articulate the balance that NWCCU strikes between autonomy and constraint:

> "Regional accreditation evaluation criteria do not prescribe specific methods nor do they dictate the characteristics and form of assessment and evaluation. Evaluation criteria do,

36

however, stress outcomes assessment as an essential part of ongoing institutional self-study, assessment, and evaluation. In accord with its mission and structure, each accredited institution is expected to formulate a plan which provides for a comprehensive assessment of outcomes and, further, to incorporate the results of assessments to improve planning that leads to successful achievement of missions and goals."

Thus, while the NWCCU standards seem less directive than some others, they contain themes common with all the other agencies: outcomes assessment is *essential*; there must be an assessment plan; and the results of assessment must be used in some way to improve.

The Western Association of Schools and Colleges, Senior College Commission

The WASC Handbook for Accreditation contains four key but interlocked standards; some even repeat passages to indicate the overlap. For example, Standard 1 "Defining Institutional Purposes and Ensuring Educational Objectives" is followed by "Achieving Educational Objectives through Core Functions." Connecting the standards through key phrases connects them as well. The Commission's holistic approach is thus different from the other commissions, and its division of its reporting/visiting cycle into two—one for determining Institutional Capacity and one for validating Educational Effectiveness—both underline the interrelatedness of the standards while emphasizing the important of examining and documenting educational effectiveness.

Thus assessment issues and strategies are threaded throughout the standards. Standard 1 links institutional purpose with educational goals or outcomes and, while the criteria do not specifically call for written program outcomes, Standard 1.2 assumes some kind of explicit statement where "Educational objectives are clearly recognized throughout the institution and are consistent with stated purposes. The institution has developed indicators and evidence to ascertain the level of achievement of its purposes and educational objectives." (2001, p. 17) Standard 2.2, on academic programs, requires that all degrees "are clearly defined in terms of entry-level requirements and in terms of levels of student achievement necessary for graduation that represent more than simply an accumulation of courses or credits." (2001, p. 20) Standard 2.4 requires that "the institution's expectations for student learning and attainment are developed and widely shared among its members (including faculty, students, staff, and where appropriate, external stakeholders)." (2001, p. 21) And Standard 4.3 incorporates assessment in planning: "Planning processes are informed by appropriately defined and analyzed quantitative and qualitative data, and include consideration of evidence of educational effectiveness, including student learning." (2001, p. 29)

WASC standards thus seem less prescriptive than some of the other standards we have reviewed, but the holistic assumption that awareness of student learning goals and the degree of their achievement are part of institutional culture and planning, means that WASC views educational effectiveness, demonstrated in part by the assessment of student learning outcomes, as a central institutional concern. Thus, while the word "assessment" may not appear with the frequency with which it appears in other accreditation agencies' statements, its presence is clearly felt throughout.

Within this holistic approach, however, WASC defines baccalaureate programs fairly specifically:

> Baccalaureate programs engage students in an integrated course of study of sufficient breadth and depth to prepare them for work, citizenship, and a fulfilling life. These programs also ensure the development of core learning abilities and competencies including, but not limited to, college-level written and oral communication; college-level quantitative skills; information literacy; and the habit of critical analysis of data and argument. In addition, baccalaureate programs actively foster an understanding of diversity; civic responsibility; the ability to work with others; and the capability to engage in lifelong learning. Baccalaureate programs also ensure breadth for all students in the areas of cultural and aesthetic, social and political, as well as scientific and technical knowledge expected of educated persons in this society. Finally, students are required to engage in an in-depth, focused, and sustained program of study as part of their baccalaureate programs (2001, p. 20).

Like several other regional accrediting agencies, WASC here defines, to some extent, the nature of general education, if not the curriculum itself. The inclusion of these outcomes, in the context of a holistic approach to assessment, suggests the assumption of broader institutional involvement in general education than is customary in American higher education.

The Higher Learning Commission, North Central Association
Like WASC, the Higher Learning Commission has a relatively small number of criteria for accreditation. These criteria, though, are refined by a series of "Core Component" statements, which articulate the general criterion, and each of which an institution is expected to address. Criterion 3: "Student Learning and Effective Teaching" shifts the traditional focus from academic programs to student learning and, in so doing, signals the assessment orientation of the criterion, the specific language of which—"The organization

provides evidence of student learning and teaching effectiveness that demonstrates it is fulfilling its educational mission"—makes the assessment emphasis explicit.

Throughout the Core Components, each of which must be addressed by the reporting institution, specific assessment themes are addressed. Core Component 3a most explicitly addresses assessment of learning outcomes: "The organization's goals for student learning outcomes are clearly stated for each educational program and make effective assessment possible." That component is followed by a series of examples of evidence, including, "Assessment of student learning provides evidence at multiple levels: course, program, and institutional," and "Results obtained through assessment of student learning are available to appropriate constituencies, including students themselves." Like the other accrediting agencies, the Higher Learning Commission does not prescribe a particular method of assessment; in fact, the furthest it goes is to suggest that assessment involves "multiple direct and indirect measures of student learning." But the emphasis on assessment as a central element of the accreditation process is clear.

That emphasis is reinforced in the other components of Criterion 3 as well. Thus, "effective learning environments" can be demonstrated by showing that "assessment results inform improvements in curriculum, pedagogy, instructional resources, and student services." Learning resources can demonstrate that it supports student learning by demonstrating that it "regularly assesses the effectiveness of its learning resources to support learning and teaching."

Even in areas outside Criterion 3, the commitment to assessment of student learning is important. Thus the resource-oriented Criterion 2 sees as appropriate evidence the provision of "adequate support for its evaluation and assessment processes," and educational quality and student learning are made explicit elements of planning.

Conclusions

All of the accreditation standards cited in this chapter are recent. Some, like the HLC criteria, went into effect in January 2005; others, like the Northeast standards, are still in draft form. The fact that none of them is more than four years old, coupled with the fact that most of them have recently given assessment a much more explicit place in the standards, suggests a very important change in higher education accreditation. What matters now is less input than outcome, less retrospective than prospective thinking, less absolute autonomy than responsibility to all stakeholders.

The variations among the regional accrediting agencies, though many, are not significant enough to suggest radically different approaches to assessment. Further, the idiosyncrasies of peer review teams mean that, even within regions, standards will not necessarily be applied uniformly at the point at which accrediting agency views of assessment are given the

Table 1
The Council of Regional Accreditation's (CREA)
Principles of Good Practice in Regional Accreditation

Area	AACSB	CRAC
Unit of analysis for assessment (Institution, college, department, degree programs, or majors)	Degree programs	Institution
Expected level of faculty involvement in assessment	High. All faculty are expected to be involved in key assessment processes.	High. Emphasis on "evidence of student learning from multiple sources" and "collection, interpretation, and use of student learning evidence [as] a collective endeavor" suggests central faculty role at all stages.
Specific guidelines for assessment of general education	No. However, AACSB expects graduates to have general skills that would normally be learned in general education courses as well as in some business courses.	None. Several describe general education outcomes in the standards, but none prescribe specific guidelines for assessing them.
Type of assessment method emphasized or required (Direct or Indirect)	Direct. Indirect may be used as a supplement.	Requires "evidence of goal attainment using appropriate assessment tools . . .from multiple sources . . . . Evidence collected from these sources is complementary . . . ." Thus both direct and indirect measures may be used, so long as they are appropriate within the context of institutional mission
Expected level of implementation at time of campus visit (or is there a deadline for mature implementation of assessment-i.e., closing the loop and creating an assessment culture)	Mature implementation expected by 2007.	No specific timeline.
Most recent update of accreditation standards	2004	2004

40

most attention—the site visit. Yet there are a number of common points about which all agencies agree, all contained in a statement of common "Principles for Good Practices in Regional Accreditation," published by the Council of Regional Accrediting Commissions and neatly summarized by Jon Wergin (2005) in the January/February issue of *Change.* As the following table shows, these common principles are relatively congruent with AACSB approaches, except in the expectation that mature assessment will be achieved by 2007 (an understandable difference since the scope of accreditation is much larger in the regional commissions) and the more permissive attitude toward indirect assessment data.

Regional accreditation, initially a device by which institutions could protect and isolate themselves, has now become a medium through which institutions explain themselves to themselves and their constituents. But those constituents have broadened; they include institutional staff, students, peer institutions, and external stakeholders like employers and donors. The use of assessment data in accreditation makes it possible for institutions to tell the world how well they are doing on what their mission statements say they value most, student learning.

Endnote
13 El-Khawas (2001) and Harcleroad (1980) trace accreditation-like processes back to 1787, when the reorganized University of the State of New York required annual reports.

References

American Council on Education, *et al.* (2004). Recommendations for the Reauthorization Of the Higher Education Act."

Baker, R. L. "Evaluating Quality and Effectiveness: Regional Accreditation Principles and Practices." *The Journal of Academic Librarianship* 28.1 (Jan-Mar. 2002): 3-7.

Commission on Colleges, Southern Association of Colleges and Schools. (2004). Principles *of accreditation: Foundations for quality enhancement.* (1st) (December 2001).

Commission on Institutions of Higher Education, Northeast Association of Schools and Colleges. (2004). *Draft Standards for Accreditation.* Version: June 2004.

El-Khawas, E. (2001) *Accreditation in the United States.* Paris: International Institute for Educational Planning.

Folger, J., & Harris, J. (1989). *Assessment in accreditation.* Washington, D.C.: Fund For the Improvement of Post-Secondary Education.

Harcleroad, F. F. (1980). *Accreditation: History, process, problems.* AAHE-ERIC Higher Education Research Report No. 6. Washington: American Association for Higher Education.

Higher Learning Commission, North Central Association of Colleges and Schools. (2004). *The handbook of accreditation,* 1(10), 3.

Middle States Commission on Higher Education. (2004). *Characteristics of Excellence in Higher Education: Standards for Accreditation.* (1st ed.) (2002).

Middle States Commission on Higher Education. 2003. "Student Learning Assessment: Options and Resources."

Murray, F.B. (2001). From consensus standards to evidence of claims: Assessment and accreditation in the case of teacher education. In *How Accreditation Influences Assessment* (J.L. Ratcliff, E.S. Lubinescu, and M.A. Gaffney, editors), San Francisco, CA: Jossey-Bass.

Northwest Commission on Colleges and Universities. *Accreditation standards and related policies.* Retrieved data, from http://www.nwccu.org/Standards%20and%20Policies/Accreditation%20Standards/Accreditation%20Standards.htm

Ratcliff, J. L., Lubinescu, E. S. & Gaffney, M. A. (Eds.). (2001). *How accreditation influences assessment.* New Directions for Higher Education 113 (Spring 2001). San Francisco: Jossey-Bass.

Wergin, J. (2005). Taking responsibility for student learning: the role of accreditation. *Change, 37*(1), pp. 30-33.

Western Association of Schools and Colleges. (2001). *WASC handbook of* Accreditation.

Author Bio

Robert Mundhenk, now Director of Assessment and Senior Scholar at the American Association for Higher Education, worked for 27 years at Northampton Community College, as a faculty member and as Vice President for Academic Affairs and Director of Assessment. In his administrative capacities he managed both regional and specialized accreditation processes, introducing assessment into each.

Recent demands for assessment of student learning haven't caught King's College by surprise — assessment of student learning has been a priority there for 20 years. King's College is one of only a handful of schools that have built a national reputation for assessment. As a result, members of their faculty are frequently requested to give presentations on King's curriculum and assessment program throughout the nation, including seven times for the American Association for Higher Education. In this chapter, King's philosophy about assessment as learning, and what it means for the McGowan School of Business, is discussed. Methods used in the Comprehensive Assessment Model – including the Competency Growth plan that establishes standards for each of the learning goals by year within the major – are illustrated.

CHAPTER 3
ASSESSMENT AS LEARNING
Cheryl O'Hara
King's College

Background, Mission, and Goals

A simple question, "What is the proper definition of excellence in higher education for students who will be living and working in the 21st Century?" was pondered at King's College more than 25 years ago. King's College continues to ponder this question, and has evolved their Comprehensive Assessment Model as a means of answering the question and continuing to strive for excellence in higher education.

Since that time, King's College's curriculum and assessment model have been featured in various national publications such as *Change: The Magazine of Higher Learning*, *The Chronicle of Higher Education* and *Barron's Best Buys in Education*. It has also been praised by national educational associations, including the Society for Values in Higher Education, and was chosen as one of only sixteen National Leadership Institutions named by the American Association of Colleges and Universities *Greater Expectations* initiative to influence the future of liberal arts higher education nationwide.

King's College was founded in 1946 by the Congregation of Holy Cross as an independent, four-year college in order to provide students with a broad based liberal education in the Catholic tradition and to offer intellectual, moral, and spiritual preparation for satisfying and purposeful lives. Education at King's College is student-centered. The faculty are committed to excellent teaching as their first responsibility, both in the core curriculum of general studies which all undergraduates follow, and in the major programs of the College of Arts and Sciences and the William G. McGowan School of

Business. The education provided at King's College promotes intellectual development, critical judgment, professional competence, reflection on religious and ethical questions, and commitment to building a just society. In the words of its founding president, King's teaches students "not only how to make a living, but how to live."[14]

King's College is located in an urban setting in Wilkes-Barre, Pennsylvania, and draws a full-time undergraduate enrollment of approximately 1,750 students, primarily from the Mid-Atlantic region and New England. The William G. McGowan School of Business enrolls approximately 450 of these students. Additionally, a limited number of part-time graduate programs are offered in specialized areas. The average class size is 20–25 students, with a student to faculty ration of 16:1. SAT scores for King's students typically range from 950 to 1,150, with an average of 1,045. King's students typically come from working-class and middle-class families, with many of them being first-generation college students. Eight-eight percent of the students receive some form of financial aid.

Before discussing the specific assessment model that King's College has developed to enable them to strive for excellence in higher education, consideration must be given to the opposing interpretations of the use of assessment in higher education. There are two conflicting perspectives on the purpose and use of assessment in higher education, the concepts of assessment as learning and assessment as measurement. In his book on assessment in education, *Enhancing Student Learning: Emphasizing Essential Competencies in Educational Programs* (1988), the late Dr. Donald W. Farmer, Vice President for Academic Affairs at King's College, differentiates between these two concepts:

> "Assessment as learning is a faculty-driven diagnostic and formative evaluation process aimed at improving individual student learning by providing continual feedback on academic performance. Assessment viewed as measuring is an administratively driven, standardized, and summative evaluation process designed to produce a numerical rating. While these two concepts are not necessarily incompatible, the primary purpose chosen will determine whether or not a college realizes the promise assessment holds for improving higher education."

The Comprehensive Assessment Program at King's College was conceived in the 1980s at a time when the curriculum for general studies was being reconceptualized. Two key elements were determined as cornerstones of the new CORE curriculum. First, an emphasis would be placed on cumulative learning, necessitating some sequencing of courses. Second, seven transferable skills of learning (Critical Thinking, Effective Writing, Effective Oral Communication, Technology Ability, Library and Information

Literacy, Quantitative Reasoning, and Moral Reasoning) would be emphasized throughout the curriculum, both in the new CORE and in curricula for individual majors. From this new CORE curriculum design, the next logical step for King's College was to consider how to assess the success of this new curriculum and use it as a means to improve the education of its students. The Comprehensive Assessment Program has evolved over time, but maintains these key elements.

With the primary purpose of academic assessment at King's College being assessment as learning, systematic feedback to students on their academic progress toward meeting the expectations of faculty throughout all four years of undergraduate studies is emphasized. For class assignments, feedback to students is usually provided by the faculty member, either in the form of written comments on the student's paper, or orally during private conferences with the student. For some assessments, other faculty members or outside evaluators are used to provide additional feedback from differing perspectives. The feedback not only points out where the student erred, but also offers suggestions for remedying the problem. In keeping with its commitment to student-centeredness, King's College maintains an Academic Skills Center which houses a tutoring program, a writing center, and a learning strategies workshop program.

The assessment program is intended to be diagnostic and supportive of student learning. Being able to act upon feedback provided by faculty members enables students to become more successful learners, and is a major feature of the model. There are multiple assessment experiences that take place for students from the point of entering King's College to the point of graduation. No one assessment experience, viewed in isolation, can ever be considered adequate.

The King's College Comprehensive Assessment Model focuses on assessing students as part of the natural teaching/learning process in the classroom. This course-embedded model does not affect the traditional definition of what it means to be a faculty member, although it does provide common assessments for students across the curriculum. Faculty members still have the autonomy to design their courses, but a minimal framework for courses in the CORE curriculum is determined by faculty members and their cohorts. Project Teams, composed of faculty currently teaching in a CORE area, convene during the semester to compare strategies and share ideas. It is during these meetings that changes can be made to the course outline. The McGowan School of Business has further developed this team approach with its common business courses. The Project Teams for these courses also include students who have completed the courses as well as members of the Business School Advisory Council.

Assessment strategies at King's College are primarily embedded in course work and therefore are a natural and integral part of the teaching/ learning process rather than being external and intrusive. They are designed

by the faculty members, not "handed down" to them by the administration. This gives the faculty members a sense of ownership not found in many other assessment models. Faculty members also are able to see the value of their work, in terms of improving student learning. They use the feedback they receive from the students to improve their courses. Thus, this assessment model is much more easily accepted by faculty members than a model with assessments that they had no voice in designing.

Assessment strategies embedded in course work also address directly the question of student motivation. Students take assessment seriously because it counts as part of the course grade, even though faculty assign to it an additional special purpose for assessing specific learning objectives.

From a faculty member's perspective, assessment as learning is inherently designed to improve teaching and learning. Any attempt to assess a specific skill or learning outcome increases understanding for both students and faculty. Used diagnostically, assessment also helps faculty members to identify students' strengths and weaknesses. This information can then be used to design appropriate teaching/learning strategies. An additional diagnostic use of assessment is to help faculty members monitor the effectiveness of the curriculum and to provide the basis for its further refinement. Assessment as learning also responds to curricular questions such as the sequencing of courses, the relationship of teaching strategies to course objectives, the responsiveness of assignments and examinations to course objectives, and whether or not the curriculum encourages cumulative learning for students.

The curriculum and assessment design of King's College's Comprehensive Assessment Model demands that faculty help students to understand the expected exit criteria for graduation and provide a plan by which students may successfully meet these standards. Assessment can contribute to student learning by encouraging faculty to make goals and criteria for judgment explicit as well as by eliciting sequential behaviors in students that contribute to their achieving the desired levels of competence. Making criteria for judgment explicit is best accomplished at the time an assignment or other assessment is assigned to the students. Clarifying the goals of an assignment or project, and providing grading rubrics, checklists, and descriptions, can aid the student in understanding what the level of expectation is prior to their starting the work. Students can't necessarily be expected to master a difficult goal with one attempt, therefore sequential behaviors to shape the desired outcome are often used throughout a curriculum. For example, an assignment in an introductory-level marketing course might require groups of students to create a simplistic marketing plan for a product. By the time students reach a senior-level marketing course, they would be required to individually create a much more sophisticated marketing plan.

From a student's perspective, the assessment model at King's College provides them with insight into what they're going to be learning in individual

courses, and indicates clearly to them the level of quality expected of them in their work. On a curricular level, this model provides the student with a visualization of the linkage between courses. This information is pointed out to the student in course syllabi, which use such phrases as "in this course you will learn …" and "as you have learned in CORE X, which is a prerequisite for this course…" Additionally, it helps them to monitor their cumulative progress, which leads them to increased levels of confidence and motivation.

The following assessments represent the major components of King's College's Comprehensive Assessment Program that occur at critical junctures in the major to ensure that students are combining learning in the CORE with learning in the major. Additionally, assessments take place prior to a student's enrollment and following graduation. Placement Tests in Critical Thinking, Effective Writing, and Quantitative Reasoning are conducted with incoming students for their assignment to appropriate level CORE courses. Students who have been accepted to King's College but are deficient in one or more of these areas are required to take remedial courses prior to enrolling in the required CORE courses. Surveys are conducted with graduates one, five, and ten years after graduation to assess the longer-term impact of their education at King's College. What all of these assessments have in common can be summarized as having:

- Clearly-defined faculty expectations for learning that students can understand;
- Explicit criteria that faculty and students can use to evaluate performance;
- Clear, honest, and timely feedback to students so they can concentrate on practical ways to improve performance;
- Strategies to enable students to connect learning in the CORE with learning in the major;
- Close collaboration and a helping relationship between faculty and students to encourage on-going development; and
- Students understanding more of what and how they learn so that they may become more involved and more responsible for their learning.

Competency Growth Plans

For each of the seven transferable skills of liberal learning, each department or program defines the skill within the context of the major and then divides the skill into specific competencies students develop from freshman through senior years in both CORE and major courses. Individuals or teams of faculty within a department originally designed these plans, which the entire department then reviewed and adopted. Each plan includes a definition of the competency, a listing of courses and assignments ("strategies") designed to help students develop the competency, and specific

criteria faculty and students use to gauge the quality of the student's performance. When read sequentially, the criteria from freshman to senior year reveal the developmental nature inherent in acquiring and promoting the skill. Faculty have found the plans very useful for tracking the transfer of liberal learning skills within the context of the major curriculum; many have remarked as well on how the plans can help departments think more carefully about curricular patterns and sequence. Rather than being given copies of these plans, students see them only in a kind of translation when objectives and criteria appear in their syllabi and instructions for assignments within courses.

The following chart illustrates a section of a Competency Growth Plan. This excerpt is from the Competency Growth Plan in Effective Writing for a student majoring in marketing. The first column ("Competency Description") explains specifically what the student should know or be able to do. The middle column ("Strategy") describes what course or courses the student will be able to develop the necessary ability and the kinds of assignments that will enable the student to achieve the competency. Finally, the last column ("Assessment Criteria") describes how the student's performance will be judged and how the instructor and student will know that the ability has been achieved.

Table 1
Sample Page: Competency Growth Plan in Effective Writing for Students Majoring in Marketing

COMPETENCY DESCRIPTION	STRATEGY	ASSESSMENT CRITERIA
	MKT 315: Consumer Behavior	
The student will be able to apply the principles of effective writing to a variety of specialized topics and audiences in marketing.	Preparation of the Sophomore/Junior Diagnostic Project	The student writes a report which is well-organized and rhetorically effective.
		The student presents material in a clear and concise manner, using charts, graphs, and matrices where appropriate.
		The student uses data and evidence of sufficient quality and quantity in support of generalizations and conclusions.
	MKT 480: Marketing Management	
The student will be able to write a formal case analysis at a professional level for a marketing management problem.	Preparation of a series of written case analyses.	The student uses the standard case analysis format to present a thorough and thoughtful solution to a case problem.
The student will be able to write a Marketing Plan at a professional level for the introduction of a new product.	Preparation of a Marketing Plan for a new product.	The student integrates relevant marketing theory to support recommendations.
		The student uses the standard format for a Marketing Plan, incorporating relevant marketing theory to support the plan.

Sophomore/Junior Diagnostic Projects

The Sophomore/Junior Diagnostic Project is designed by faculty in each major program for their majors. The assessment is designed to serve as a diagnostic screening device to determine the ability of students to transfer learning and skills developed through

the CORE curriculum to a selected question, case study, or project related to their major field of study. This assessment takes many different forms, as dictated by the faculty of each major. The design can also take many different forms, such as a research memorandum, portfolio, written report, or formal presentation. The Sophomore/Junior Diagnostic Project provides a process check for Competency Growth Plans which integrate learning in the CORE curriculum within the context of the major program.

The project is typically evaluated by the department faculty, although outside evaluators are used in some departments. Feedback is provided to students by either the instructor in the designated course or by the individual student's academic advisor. The Sophomore/Junior Diagnostic Project helps students to develop a clearer understanding of the expectations of faculty in their major field of study, with respect to their ability to apply critical thinking skills and to communicate effectively. It also helps students to develop a better understanding of the specific criteria that faculty use to judge work of students in their respective major fields of study. For those students who do not reach the level of competency expected, remedial work is usually prescribed. Often, this is coordinated with the help of the college's Academic Skills Center. The student must then re-take the assessment to show that they have mastered the work. Failure to successfully complete the assessment also leads to a discussion between the student and his or her faculty advisor, to ascertain if the student should continue to pursue the major.

Senior Integrated Assessments

The culminating assessment experience for students at King's College is the Senior Integrated Assessment. This assessment is intended to provide an opportunity for the faculty in a student's major field of study to make a holistic judgment of the student's education, especially the ability of the student to integrate the transferable skills of liberal learning with learning in his or her major field of study. As with the other assessments in King's College's Comprehensive Assessment Program, this assessment is course-embedded. In creating these assessments, most departments chose to integrate them into existing capstone courses for the major.

The integrated assessment is intended to be a performance-based student experience that can be evaluated by all faculty members in the student's major field of study as well as by professionals in related fields of employment, if appropriate. The assessment experience provides the basis for evaluating the following areas of learning:

- Command of the knowledge base for the major field of study;
- Mastery of the methodology of the major discipline; and
- Competence in the transferable skills of liberal learning relating to the departmental competence statements and four-year competency growth plans.

49

Within this framework, each department creates a meaningful assessment that is relevant to the major. This can take many forms, such as research projects, laboratory experiments in the sciences, case studies, or portfolios.

Feedback is not only important for students, but also for the faculty. The results of the various assessments are collectively used by the faculty to make any necessary adjustments or changes in the curriculum. The business school annually conducts an "Assessment Day" at the end of the spring semester. At this time, faculty members from each department report the results of their assessments and discuss what changes, if any, need to be made for the following year. This information sharing often results in a lively dialog between faculty members in different disciplines and improve the ability of faculty members to see the effects that changes to their curricula might have on other departments.

Conclusions

As shown, the Comprehensive Assessment Program at King's College has the primary purpose of enhancing student learning. As has been shown, the model designed and used at King's College clearly gives ample opportunity for students to monitor their progress and improve their academic performance. However, this model of assessment does not preclude using it for assessment as measurement, which is often required by outside accreditors and other external stakeholders. King's College has successfully used this model for external validity of their level of excellence in higher education. Primarily, sampling is used to demonstrate student success at achieving the predetermined goals that have been set for each skill area and for each of the majors. Thus, the two interpretations of assessment—assessment as learning and assessment as measurement—can co-exist.

Endnote
14 Excerpts from the King 's College Mission Statement

Author Bio

Cheryl O'Hara joined the marketing faculty of King's College in 1979 and served as the Chairperson of the Marketing Department from 1983 to 1996. Currently, she holds the position of Associate Director for Continuous Improvement and is responsible for the assessment program within the business school. For many years she has been very active in the design and implementation of the assessment model used at King's College. She makes frequent presentations on it, and has also served as a consultant on assessment to many colleges and universities nationwide.

SIUE is another "old hand" at assessment of student learning, and there are many lessons that can be learned from their journey. Implemented in the late 1980's as part of the university's articulated priority on teaching and the adoption of the Scholarship of Teaching model to document teaching excellence, assessment became a way of life for faculty there. Supporting mechanisms including staff support, faculty development, and permission to fail, created a campus culture in which assessment is valued, not feared.

CHAPTER 4
A CULTURE OF ASSESSMENT

Douglas J. Eder
Southern Illinois University Edwardsville

Background, Mission, and Goals

Southern Illinois University Edwardsville (SIUE) began operating its assessment program in the Spring of 1989. Through a combination of listening, experimenting, failing, persisting, and rewarding, SIUE created a history and a campus culture that embraces assessment. The embrace is less than passionate but much more than perfunctory. This culture came into being because of early, intentional, and periodic interactions with several major players in the assessment movement. By listening carefully to those players and heeding their advice, SIUE avoided most of the caustic arguments and entropic errors that characteristically hindered progress elsewhere. This allowed the university to build an identifiable and positive culture of assessment. Building that culture was not an easy, accidental, single event. While avoiding many common entanglements, SIUE insisted on ensnaring itself in one major web of its own making. Moreover, circumstances evolved as veterans retired and new actors entered, thereby causing some previously solved problems to resurface. Nevertheless, the campus outlook on assessment is positive, and the university has used assessment to improve its students' learning and its faculty's teaching practices. We believe that some of the lessons learned are generalizable and, therefore, we recite them here.

The SIUE mission statement describes the university as a public, comprehensive institution dedicated to the communication, expansion, and integration of knowledge through its first priority, excellent undergraduate education. The presence of excellent graduate and professional academic programs complement this mission. SIUE was chartered in 1957 as an outgrowth of SIU Carbondale. It became an independent campus with its own physical facilities and regional mission in 1965. The full-time instructional faculty numbers more than 500. Student enrollment is now approximately

10,500 undergraduate and 2,500 graduate students. Almost 25% of the student population lives in campus housing.

A Wee Bit of Context

During the mid-1980s national public attention focused visibly on the condition of U.S. higher education. Reports appeared such as *A Nation at Risk* (National Commission on Excellence in Education, 1983), *Involvement in Learning* (National Institute of Education, 1984), *Time for Results* (National Governor's Association, 1986), and *Seven Principles for Good Practice in Undergraduate Education* (American Association for Higher Education, the Education Commission of the States, and the Johnson Foundation, 1987). These reports asked such questions as:

- Are college students learning what we say they are learning?
- Are college students learning what they <u>ought</u> to be learning?
- Can college students communicate well and solve problems effectively?
- Are college graduates able to apply their knowledge and skills in the workplace and/or in advanced educational environments?

Only a few institutions responded to these questions by publicly re-examining their own accountability for academic quality. SIUE's response became visible by 1989 with the activation of its Assessment Plan. Assessment's origins at SIUE derive primarily from two sources: (a) a history of state-driven program reviews and cost studies extending back through the 1970s, and (b) the premonitions of SIUE's then-President Earl Lazerson, whose prescient survey of the mid-1980s scholastic landscape suggested correctly that calls for educational accountability were enduring and would not go away. In a two-year-long response to administrative exhortations, a task force composed of faculty, administrators, staff members, and students constructed the SIUE Assessment Plan. In the fall of 1988, U.S. Secretary of Education William Bennett directed that all federally approved accrediting organizations include evidence from outcomes assessment in their (re-) accreditation standards. When this finally caught the attention of higher education in general, SIUE was already poised to embrace change and assessment due to the astute thinking by senior members of its faculty and administration. In addition, the university culture was starting to convert from one populated by its founders to one run by its next generation.

> **Lesson:** Perceptive university leadership helped SIUE see assessment as the arrival of an enduring call for educational accountability rather than the temporary interference of a passing fad.

During the early 1990s, SIUE undertook a new look at itself while simultaneously completing a structural reorganization that created a College of Arts and Sciences, converting its academic calendar from quarters to semesters, and completing the step-wise installation of its assessment program. Coincidentally, the university searched for and acquired a new president in Nancy Belck who, along with Provost David Werner, sensed the need to involve the entire university community in reconstructing its mission, goals, vision, and values. There was a simultaneous need to reform tenure/ promotion guidelines and to align the budget with rejuvenated aspirations even as the university hired new faculty members to fill the vacuum of retiring veterans. Over the course of about four years, SIUE engaged in several "advances" (as opposed to "retreats"), during which all constituency groups proposed, argued, negotiated, and (mostly) resolved the university's big issues. Results included creation of a concise mission statement, a set of briefly and clearly stated university goals, agreed-upon vision and values, a new policy for faculty tenure and promotion, department-by-department learning objectives for students, and some mechanisms by which to monitor progress.

Doing all this simultaneously was not easy, and multiple factors contributed to concurrent progress on all items. One key contributor to successful change was the openness of the Calendar Conversion Committee, especially its faculty chairperson, whose quarter-to-semester conversion handbook and frequent, detailed minutes modeled how to invite faculty participation and announce results. Another key contributor was the initiation of an extensive faculty development program. Faculty members were beginning to do things they had never done before. They asked for, and received, high quality help in the form of seminars, workshops, and trips to national forums on learning, teaching, and assessing. Visitors to campus included assessment innovators and early participants in the Scholarship of Teaching thinking as envisioned by the Carnegie Foundation and the American Association for Higher Education. For example, some of our early visitors were Pat Hutchings, Dan Bernstein, Barbara Cambridge, Charles Glassick, and Peter Shedd. Faculty sojourns elsewhere included team attendance at various American Association for Higher Education (AAHE) forums, the American Association of Colleges and Universities (AAC&U, at that time simply the AAC) national conferences, and the University of Prince Edward Island's (UPEI's) Teaching in the Active Voice Institute. Participants in these sojourns were subsequently tapped for on-campus seminars and other roles in leadership and facilitation.

A third contributor to progress was the use of "advances" for the purpose of creating a culture of discussion, and their performance merits description. In general, each advance lasted 3-7 hours, was led by one or more capable and respected facilitators, and involved an invited group of perhaps 75 diverse individuals selected from all constituencies. On the appointed day, individuals distributed themselves to assigned tables, each of which had a table flag that

identified a particular role. For example, construction of our university statement of values began, due to local history, with broad consensus around respect for cultural diversity, service to the region, civic engagement and leadership, a sense of environmental stewardship, and a yearning to promote life-long learning. Therefore, printed flags with these sorts of titles already adorned many tables. Blank table flags were available to accommodate those individuals who felt strongly about discussing additional items. Each table's role was to construct a concise text around its specific topic. After an hour of table discussions, each group wrote its text on a flip chart and reported to the whole assembly. Lunch ensued, a very important component of the advance because it promoted collegiality. Meanwhile, the facilitator(s) collected the flip chart pages, which were transcribed and edited into a coherent working document in the president's office. After 4-8 weeks, the working document was circulated as a draft throughout the university community via e-mail. Subsequently, the 75 participants reassembled for 2-3 hours to address the question, "Does this document reflect what we really mean?" Once again, the facilitator(s) collected flip chart pages and the president's office edited the revisions prior to sending a final document on to the Deans' Council and the Faculty Senate for formal consideration. Naturally, because numerous constituencies had their fingerprints on the document, acceptance and passage of new policy was relatively speedy, the entire gestation from concept to policy taking place in one academic year. Together, a combination of open practices, faculty development, and the spirit of "advances" generated the critical mass necessary to sustain a campus-wide momentum and culture for change.

> **Lesson:** Administrative leaders negotiated clear goals that flowed logically from our history and circumstances.

> **Lesson:** We invited experts to help us, we listened—really listened—to their advice, and then we took it.

> **Lesson:** New faculty roles and expectations were accompanied by visible and high quality faculty development.

> **Lesson:** Faculty-administrative collaboration on new policies was real, and it was perceived as real. The resulting policy changes were communicated widely and openly.

Faculty Ownership of the Jewel: SIUE's Senior Assignment

SIUE's Assessment Plan established three times in a student's career for assessment: upon entry, at midpoint, and during the senior year. Entry assessment was designed to help characterize the kinds of students who matriculate as first year and as transfer students. This kind of assessment

involved standardized tests, placement tests, and surveys. Midpoint assessment involved a Rising Junior Paper (*vide infra*). Senior year assessment featured what has become the jewel of SIUE's assessment program, the Senior Assignment (SRA). The Senior Assignment is defined as a scholarly, (semi-) independent student project that is conducted under the supervision of a dedicated professor and that results in a visible product or behavior. As such, student learning, and the curriculum that produced it, can be assessed. SIUE professors *as a faculty* care about the Senior Assignment concept, because early on they took the position that students should not be able to earn a college degree merely by being adequate classroom stenographers. Rather, candidates for graduation should have taken control of their own education and actually *done* something with it. Therefore, the Assessment Plan placed the SRA directly into the hands of the faculty by giving it departmental ownership, subject to the guidelines of the Plan. The Plan placed assessment operations under the responsibility of a supervisory Committee on Assessment (COA) and its full-time Director of Assessment, the latter being a faculty member who is paid by the provost and who reports to the provost *and* to the Faculty Senate. This arrangement accomplished three things: (a) It gave to the faculty primary control over something it cared most about—the SRA— and it charged to an administrative office those things the faculty had less patience for—e.g., administering placement tests and conducting surveys; (b) it endowed assessment with visibility in the provost's office while simultaneously maintaining ownership by the faculty; and (c) it identified an individual director as responsible for assessment operations, thereby avoiding the "Fallacy of the Commons", where everybody is supposed to be responsible but, in actuality, no one is.

> **Lesson:** Faculty members assessed the things that mattered most, unhampered by interference from the things that mattered least.

> **Lesson:** Responsibility for assessment success was assigned to a visible, identified individual with direct access to university administrators and a clear chain of accountability to the faculty.

Installation: Permission to Fail but Not to Stall

In the late 1980s, three major changes were beginning at SIUE: Planning to integrate nine schools into one College of Arts and Sciences and five schools, planning for calendar conversion from academic quarters to semesters, and installing academic assessment. In order to acquire faculty buy-in, the Committee on Assessment installed assessment in stages that tracked the entering class in Fall 1989. Accordingly, placement tests for entering students began in 1989, but the Senior Assignment did not begin

55

until 1992. This arrangement caused the entering class to experience a fully operational, four-year assessment plan as it moved through, even while the faculty experienced a four-year phase-in. During 1992-93, coincident with assessment's fourth year of operation, the university completed its reorganization from nine schools to a College of Arts and Sciences and five professional schools. It also converted its academic calendar from quarters to semesters. Given the convergence in 1992 of three simultaneous major changes in the university, many departments appealed to COA to delay implementation of the Senior Assignment. The provost saw that implementing the Senior Assignment in a timely manner was the right thing to do, even if initially it was done imperfectly. Therefore, appeals to prolong the process of installing assessment were denied. Thus, by 1992, the Assessment Plan was fully operating and doing so in a new academic environment.

"Institutional effectiveness," said Peter Ewell (1985), "is a comparison of results achieved to goals intended." Because SIUE's Senior Assignment would become a reality for the first time in 1992-93, the Committee on Assessment asked each department in 1989 to begin establishing goals and objectives for student learning in its programs. This consideration of goals and objectives coincided with the revamping of the curriculum for calendar conversion. As part of the general reorganization, the new goals and objectives could be discipline specific but, in some way, had to embrace university-wide general education objectives. Similarly, the venues in which assessment would occur were up to each department as long as each departmental faculty as a whole examined its students' products and witnessed its students' behaviors. The process of establishing initial goals and objectives and matching appropriate assessment mechanisms to them was a challenging one, one that lasted into 1992.

Here is an example that illustrates the nature of the challenge and the kind of response that followed in order to induce faculty ownership. One department's list of student learning goals included the statement that students "should have the oral competency to deliver a talk on a technical topic to a lay audience by using jargon-free speech." This same department proposed using a nationally normed, standardized, written exam as its assessment device. Upon receiving this department's assessment plan, the University Committee on Assessment muttered some derogatory internal comments like, "Department X is (a) outsourcing its responsibility for assuring student learning, (b) invoking a written assessment device to assess an oral competence objective, (c) misusing the entire concept of standardized testing by failing to appreciate the statistical basis of 'norming,' and (d) demeaning the spirit of improvement around here." All true. What COA did publicly, however, was to engage the department in a patient, sincere conversation that went something like this: "We appreciate your initial, rapid response to our request. Our understanding of your discipline is incomplete and we need some help. Would you please explain how a written test that is taken nationally by students

motivated to enter graduate school will reveal, to your faculty's satisfaction, your undergraduate students' achievement in oral communication? We are sure you have thought this through. Is it possible that we have not received all the pages of your assessment proposal?" Faced with a sincere, gentle, yet direct critique of its plan, the department reflected for a month and produced an improved, reasoned plan that satisfied COA's minimum criteria and matched assessment methods to goals and objectives.

Word of encounters like this one spread quietly and helped to reinforce the notion that assessment was less of a threat than first feared. Moreover, COA internalized the advice of wise counselors, who predicted that imperfect plans would improve with experience and time. Several departments soon became impatient with the less-than-adequate performance of their students and began to revise and align assessment tasks with classroom pedagogy. Indeed, this happened through a process of continuous improvement that was reinforced by public faculty development seminars and private individual consultations. Other departments saw assessment and its accompanying Senior Assignment Fund (*vide infra*) as opportunities for experimentation and creative adventure. Some individuals in these departments became not just stewards of assessment but outright champions. They sensed an opportunity to experiment without fear of punishment if they failed. They also saw tangible support for experiments that allowed students to achieve—and to demonstrate that they had achieved—improved learning. Celebration of success prompted emulation by other departments. The creativity generated by our pioneering champions has proved infectious, and the payoff to the university has been incalculable.

> **Lesson:** Implementation was staged over a clearly understood, multi-year time line. This time line permitted experimentation but it did not permit outright procrastination.

> **Lesson:** We gave permission to try and fail. If one is doing the right thing, one improves through practice. If one is doing the wrong thing, even if it's done perfectly, it's still the wrong thing.

> **Lesson:** We identified champions as well as stewards. Stewards take care of things, but champions, given permission and encouragement, make things happen.

> **Lesson:** Practice and improvement required time and feedback. Colleges and universities move slowly but, given appropriate and continuous stimulation, they do move.

Money
When the Faculty Senate received the Assessment Plan, a crucial

paragraph was appended due to the convictions of one particularly wise faculty senator. The key sentence of this paragraph is: "Should the state neglect its responsibility for funding expansion of assessment activities at SIUE, no internal reallocation of funds from academic units or academic support units shall occur for assessment purposes." The Assessment Plan, with this key sentence intact, was accepted and signed by the university administration. In short, when other universities were digging in their heels and being dragged toward the cliff edge of external accountability, SIUE practiced some astute critical thinking and decided to go along sincerely and openly willing…and with its palm outstretched. As a consequence, both the state itself and the university's accreditors accepted SIUE as an assessment innovator. The former added money to the university's base budget, and the latter offered latitude in reaccreditation processes in order to allow for innovation.

Part of the new money received from the state became a $120,000 Senior Assignment Fund to enhance departmental assessment. The 2006 academic year is the 12th year of the Fund's existence, which was established specifically "to foster closer student-faculty academic relationships," such as co-authored presentations, publications, explorations, and commensurate activities in the performing and studio arts. Significantly, the Fund has two components: The larger component supports major departmental enterprises, some of which may cost up to $15,000. In general, it is available to any one department once every three years. The smaller component is available yearly on an *ad hoc* basis to support ongoing activities that result from major initiatives. This means that every department can receive Senior Assignment support every year, and major support every third year. In practice, the Fund has never run dry, and some departments have received assessment funding of more than $10,000 *every* year. In addition, a separately funded and competitive Undergraduate Research Academy supports individual student Senior Assignments at the honors level. Thus, the existence of regularly available money dedicated to enhancing departmental assessment operations and student achievement has augmented a visible culture of assessment at SIUE.

> **Lesson:** Going along early and willingly with a hand out was more productive than just going along willingly, or not going along at all.
>
> **Lesson:** As the saying goes, one can lead, follow, or get out of the way. Once it was decided that assessment could be morphed into opportunity, there was more freedom in leading (and setting our own agenda) than in following (and having to follow agendas set by others).
>
> **Lesson:** Money for assessment came from new sources

and did not squeeze existing department or school budgets. In fact, assessment added money to schools and departments.

One Very Large Pothole: The Rising Junior Paper

The single component that was most toxic to SIUE's culture of assessment was the Rising Junior Paper (RJP). As the second part of the three-part Assessment Plan (assessment at student entry, midpoint, and senior year), the RJP was designed as an earnest attempt to focus on and improve student writing. The Assessment Plan specifically required "rising juniors" (second semester sophomores and first semester juniors) to take a writing flag course, that is, one that had been approved by the Committee on Assessment as containing a suitable, significant writing assignment. Each academic term, some 700 papers from these courses were collected, scrubbed of all faculty comments, grades, and identifying marks, duplicated, and then presented on a Saturday morning to a panel of about 25 faculty readers who had been trained for this task. Each reader read at least 40 (some read many more) papers during a reading session that lasted from roughly 9 am to 3 pm and received $250 for the work.

Each paper was assessed on the basis of five criteria: match to the assignment, coherence, use of evidence, mechanics of writing, and bibliographic support. A reader assigned a single, holistic score to the paper based on a 4-point scale, a value of 4 or 3 representing "above competence" and a value of 2 or 1 representing "below competence." Two readers independently read each paper, and if the readers disagreed across the line of competence, a third reader rendered a deciding score. After a reading session, the Director of Assessment processed the scores and returned them to the professors from whose courses they came and also to each student's academic advisor. Because the RJPs were collected at the end of a term, the assessment scores were never available to the professors, advisors, or students until midway through the next term. For the Spring term, this wait lasted from May to October, an interval of five months.

The RJP derived from a model used by the Educational Testing Service (ETS) for reading and scoring student essays in a standardized manner. Several factors of this otherwise successful model worked together against success at SIUE. First and most important, individual professors felt they were being second guessed, especially if "below competence" scores were returned on papers that a professor had graded as, say, B+ or A. Second, the criteria used for RJP assessment were out of register with those used by professors to grade the paper in the first place. Professors' criteria were mostly driven by disciplinary content instead of writing fundamentals. Third, papers received holistic scores, not parsed scores that would have shown exactly where efforts at improvement should be focused. Fourth, feedback, such as it was, arrived too late to have any effect other than to provoke

irritation. No specific consequences or recommendations flowed from the scores. Fifth, several operational factors that are crucial for ETS success with this model were omitted from SIUE's practices. Two of these factors were the presence of table supervisors and a workflow that included previously scored, but clandestinely presented, dummy papers. At ETS, these factors help to keep individual readers from drifting off calibration. They were not present at RJP reading sessions.

During the span of five years, hate mail spawned by the RJP poisoned assessment relationships and paralyzed new initiatives. At the time, COA was split on whether to eliminate the RJP, given its presence as policy in the official Assessment Plan. External stimulus arrived from the new university president in the form of the statement, "I'm used to a more active assessment program." Consequently, the assessment office invited a select group of regular readers to re-read RJPs one week following an official session. The group was picked to include key, vocal advocates, and even zealots, who favored the RJP, experienced readers whose judgments were infrequently overturned, and readers who were known for their integrity. This group re-read approximately 100 papers, including some that had received a pair of 4s and some that had received a pair of 1s. Some papers were routed to new readers and some were routed back to readers who read them the first time. Additionally, readers were asked to track the contribution of the five criteria to their holistic score and to record the clock time at which they read each paper.

The results were revealing.

(1) Scoring could not be distinguished from random. That is, the probability of a paper receiving any combination of two holistic scores could not be statistically distinguished from what would have happened through applying scores by flipping coins.

(2) Overall, the probability of a paper being reversed in terms of its assessed level of writing competence was 30%. This probability held whether a paper was re-read by a new pair of readers or, amazingly, by the original pair.

(3) On an individual level, the readers most likely to reverse an assessment of competence included those readers who most zealously favored the RJP and their own participation in it.

(4) Finally, with the passage of time (and onset of fatigue), holistic scores more closely tracked partial scores on writing mechanics to the exclusion of the four other criteria.

Armed with these data, the Faculty Senate solicited revisions to the Assessment Plan. COA offered language that suggested an unspecified "midpoint assessment" (e.g., portfolios, readiness tests, reflective essays, interviews) rather than a specific recipe, such as had existed for the RJP. The removal of process statements from the policy document allowed for flexibility

and evolution of assessment. Meanwhile, COA announced that five years of RJP experience showed that roughly one half of the university's juniors couldn't write up to university standards and that individual departments should take appropriate (but unspecified) action. With this, the RJP died to the collective relief of the university community. During the next two years COA membership changed, and the assessment office promoted a low profile while rebuilding relationships. To this day, it is not possible to find a single individual who advocated the original RJP idea.

> **Lesson:** The Assessment Plan, which is a policy document, prescribed a particular assessment process. Because the process was locked into policy, the only way to change the process was to amend the policy, a very cumbersome method that required a vote by the entire Faculty Senate.

> **Lesson:** Sometimes simply declaring victory and moving on is the best thing to do.

> **Lesson:** Good judgments came from experience, and much of that came from bad judgments.

The Scholarship of Teaching: One Key to a Culture of Assessment

Many institutions expect their faculties to act as institutional "good citizens," but they do not align their reward systems to honor this behavior. Professors receive greater external and internal rewards for publishing than they do for directly improving the learning of their own university's students. Thus assessment, which has both an external mandate and a potential for producing marvelous student learning benefits, fails to acquire faculty ownership. As dust settled from structural and curriculum revisions, the SIUE faculty and administration took on the issue of aligning faculty roles and rewards with the university's mission. By the Fall of 1994, SIUE's new tenure and promotion policies recognized three performance levels (satisfactory, meritorious, excellent) in each of the three traditional areas of faculty performance (teaching, scholarship, and service). For tenure and promotion, faculty members must now display at least satisfactory performance in one area and at least meritorious performance in the other two, one of which *must* be teaching. In other words, merely adequate teaching will not contribute to tenure or promotion, regardless of a faculty member's merit in the other two categories.

An institution that requires an elevated level of performance in teaching is morally obligated to provide guidance on how to reach that level. The unifying principle behind SIUE's guidance for its faculty came partly from Ernest Boyer's book, *Scholarship Reconsidered*, and the movement that it produced, subsequently known as the Scholarship of Teaching. At its core, teaching is

more than a series of scholarly and interesting classroom performances. Teaching implies student learning. Under the Scholarship of Teaching, a tenure and promotion dossier should make a case that is objective, public, reflective, and peer reviewable. Moreover, the case requires evidence in support of the claim that students actually learn as the result of the professor's efforts. A large fraction of such evidence arrives through the results of assessment.

SIUE's invocation of the Scholarship of Teaching has brought assessment out of the realm of mere data collection and into the realm of scholarship. Faculty ownership of assessment fortifies tenure and promotion endeavors. This role of assessment continues to attract reinforcement through faculty development visits by scholars of teaching and learning, including Dan Bernstein (again!), Trudy Banta, Carol Geary Schneider, Lee Shulman, and Mary Taylor Huber. At its best, faculty ownership of assessment, from classroom through program levels, gives validity to the claim that our students are learning. The payoff to the university is large.

Partly because of its culture of assessment, SIUE is a charter member of AQIP, the Academic Quality Improvement Program of the North Central Association's Higher Learning Commission. SIUE's standards of quality and its decision-making mechanisms are public. As a member of AQIP, SIUE is exempt from the onerous and expensive reaccreditation self-study and team visit. Instead, the university establishes its own 3-5 institutional questions to pursue and engages in objective, public, reflective, and peer reviewable searches for answers.

Assessment synergy has germinated between academic departments and the university as well, as shown by two closing examples. In 1996, liberated from the RJP and flush with Senior Assignment success, COA invited departments to generate their own, independent assessment projects under the banner of PILAF (Program-Initiated Learning Assessment Fund—COA enjoyed using food concepts as inducements for assessment activities). Figuring on a modest response, the assessment office promised $1,500 of unrestricted support to each department that proposed and undertook a project. Surprisingly, fully one-half of all university departments participated, including *all six* departments in the School of Business. As a consequence, business added $9,000 to its 1996-97 budget and produced six *extra* assessment reports that COA did not have to organize. Moreover, the university was able to showcase 18 diverse assessment projects, all of which arose voluntarily from within departments themselves.

The final example describes the contribution of our Assessment Scholars to a campus culture of assessment. Concurrent with the demise of the Rising Junior Paper, assessment began its Wing Portfolios (so named because it started in a residence hall wing). Students volunteered to collect specified items into portfolios and handed in their portfolios at the end of each semester, in exchange for which they received university logo clothing. The specified portfolio items were: (a) all syllabi, (b) all written papers that

students had received back, (c) all exams that they had received back, (d) an attitude survey or study log, and (e) a brief reflective essay on a topic assigned by the assessment office. The assessment office photocopied the portfolios and returned originals to the students. (*Nota bene*: University Counsel cleared all procedures, and students signed informed consent contracts in order to participate.) The Assessment Office followed six consecutive years of entering classes and collected approximately 300 portfolios. The collection provides a picture of the SIUE curriculum. Demographics of the participants shadow those of the student population in key factors except gender; females participated in the Wing Assessment to a greater extent than they are represented in the SIUE student population.

The purpose of the Wing Portfolio is to assess curricular performance, not student performance. For instance, student writing has historically concerned the SIUE faculty. How many writing assignments does, say, the average first-semester sophomore undertake? One can simply count the assignments that are represented throughout the Wing Portfolios. Or, what *kinds* of writing do the typical second-semester juniors do across all disciplines? Analytical? Reactions to a text? Comparison-contrast? Creative? Journal? One can analyze student writing within the Wing Portfolios and count the number of papers in each category.

On a more sophisticated level, assessment asked the question, "To what extent and by what means do students improve their writing at this university?" To approach an answer, COA sought three faculty scholars to examine independently the Wing Portfolios as primary literature. Three scholars, one each in speech communication, philosophy, and educational leadership, accepted the challenge. During the next year and a half, each scholar approached the task from an individual, disciplinary viewpoint. Assessment paid each scholar a significant honorarium, one third of it up front and the remaining two thirds upon receiving a report of the findings in the form of a manuscript ready for submission to a refereed journal, complete with stamped, addressed envelope. All scholars fulfilled their tasks; one manuscript has been published and one remains under review.

All three scholars presented their work publicly at university forums. Their findings are significant. One scholar compared standards and pedagogies between SIUE and other universities. Another scholar analyzed the longitudinal effects of writing practices within SIUE. The third examined the kinds of feedback that professors provide on student papers, that is, the notes to students that professors write in the margins as they read. The most common kind of feedback was a trail of marginal annotations on the mechanics of writing (grammar, word choice, missing words). Uncommon was constructive feedback about such things as the structure of an argument or the use of evidence to support a claim. Relevant feedback is important for improved performance. What kinds of messages are professors sending to students and, given their connection to grades, how do these message encourage (or

fail to encourage) the longed-for improvements in writing? SIUE's Assessment Scholars have uncovered important information about how we teach and appraise student writing, information that should lead directly to improved student learning. They did this as practitioners of the reflective Scholarship of Teaching, supported by a campus culture of assessment.

> **Lesson:** Aligning faculty roles and rewards, including financial rewards, under an umbrella of sincere scholarship has produced magnificent payoffs.

Item Arising

Assessment at SIUE remains a mobile concept. It is not a finished product, and its operation demands continuous attention and adjustments. Time has transformed the findings of the assessment scholars from memory into history. Several other solved problems have exhibited disturbing resurrections, not because of wicked intent but, rather, merely because new faculty members, staff members, chairpersons, deans, and administrators replace acculturated veterans of those same positions. Even the temporary absence of a champion permits cultural drift to occur. Thus, it has been necessary to return to old issues and to resolve them in the presence of new audiences. The energy occupied in keeping old problems solved takes away from the energy needed to solve new problems. Nevertheless, assessment movement exists and, as with a battleship, even slow, steady, forward progress conveys undeniable, maybe irresistible, momentum.

> **Lesson:** When you're riding ahead of the herd, it's a good idea to take a look back every now and then to make sure it's still there.

References

Boyer, E.L. (1990). *Scholarship reconsidered: Priorities of the professoriate.* San Francisco: Jossey-Bass.

Chickering, A.W., & Gamson, Z.F. (1987). *Seven principles for good practice in undergraduate education.* AAHE Bulletin (March).

Eder, D.J. (2004). *General education assessment within the disciplines.* Journal of General Education 53(2), 135-57.

Ewell, P. (Ed.). (1985). Assessing educational outcomes. *New Directions for Institutional Research, No. 47.* San Francisco: Jossey-Bass.

Ewell, P.T., & Jones, D.P. (1996). *Indicators of "good practice" in undergraduate education: A handbook for development and implementation.* Boulder, CO: National Center for Higher Education Management Systems (NCHEMS).

Halpern, D. (2002). *The engaged learner: Why learning is not a spectator*

sport. American Association for Higher Education Annual Conference on Faculty Roles and Rewards, Phoenix, January.

Halpern, D.F. & Hakel, M.D. (2003). Applying the science of learning to the university and beyond: Teaching for long-term retention and transfer. *Change* (July/August), pp. 36-41.

Marchese, T. (2001). *Brain research and the faculty role.* American Association for Higher Education Annual Conference on Faculty Roles and Rewards, Tampa, February.

National Commission of Excellence in Education. (1983). *A nation at risk: The imperative for educational reform.* Washington, DC: U.S. Government Printing Office.

National Research Council. (2000). *How people learn.* Washington DC: National Academy Press.

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

Task Force on College Quality, (1986). *Time for results: The governors' 1991 report on education.* Washington D.C.: National Governors' Association Center for Policy Analysis and Research.

Walvoord, B.E. & Anderson, V. (1998). *Effective grading: A tool for learning and assessment.* San Francisco: Jossey-Bass.

Weinstein, L. (1993). *Moving a battleship with your bare hands: Governing a university system.* Madison, WI: Magna Publications.

Wiggins, G. (1998). *Educative Assessment.* San Francisco: Jossey-Bass.

Author Bio

Douglas Eder is a member of the biology faculty at Southern Illinois University Edwardsville (SIUE), specializing in neuroscience. He has been steward of SIUE's assessment program since 1993 and was named Visiting Distinguished Scholar at Hamilton College in 2000-01. He publishes in both his discipline and in academic assessment, and co-presents the AACSB Assurance of Learning seminars with Kathryn Martell.

The Monfort College of Business (MCB) at the University of Northern Colorado is the only business school in the country to earn the prestigious Malcolm Baldrige National Quality Award (MBNQA). A relentless drive for excellence in undergraduate education — reflected in the college's mission and supported by its strategy ("high touch, wide tech, and professional depth"), faculty and resource deployment — yielded tremendous dividends and established the college as a leader in quality business education. The college's ability to document how its curriculum, processes and environment led to significant improvements in student learning was critical to the successful application for the Baldrige award. Thus, assessment of student learning was key to demonstrating the college's commitment to the continuous improvement that the MBNQA celebrates.

CHAPTER 5
PURSUING A QUALITY-BASED STRATEGY: A CASE STUDY OF THE FIRST BUSINESS SCHOOL TO EARN THE BALDRIGE AWARD

Timothy E. Jares
Joe F. Alexander
Kenneth W. Monfort College of Business, University of Northern Colorado

Background, Mission, and Goals

The University of Northern Colorado's (UNC) College of Business was established in 1968, with a primary mission to provide graduate and undergraduate business education. The college's evolution through the 1970s paralleled the business school national trend of explosive enrollment growth and program proliferation. By 1984, the college's 50-person faculty was serving more than 2,000 students enrolled in a wide range of undergraduate, master's, and doctoral degree programs.

A Quality Journey Begins

In 1984, the College took dramatic steps to make program quality its top priority. At the time, UNC's business program was generally regarded as average and largely overshadowed by its regional peers. With most U.S. business programs opting for a growth strategy of degree program assortment and further proliferation of graduate programs, UNC's business administrators and faculty chose a counter approach. A vision was cast for becoming Colorado's best *undergraduate* business program—a goal it was agreed would

Figure 1
A Singular Focus on Undergraduate
Business Excellence (2004)

MCB's Mission
Our mission is to deliver excellent undergraduate business programs that prepare students for successful careers and responsible leadership in business.

MCB's Vision
Our vision is to build a reputation of excellence in Colorado and beyond for preparing future business leaders and professionals.

MCB's Values
Each MCB value statement is held within an overall framework focused on the pursuit of excellence; a philosophy of continuous improvement guides employee behavior.

only be possible by making undergraduate business education the college's exclusive mission (see Figure 1). Within two years, a revolutionary plan commenced for eliminating all graduate programs, including a PhD and large MBA program. Changes also were made at the undergraduate level, with the elimination of all but one degree program—the Bachelor of Science in Business Administration. Future UNC business students would enroll in a single business major and choose from six emphasis areas: accounting, computer information systems, finance, management, marketing, or general business. The college adopted two long-term strategies to guide its actions: (a) a program delivery framework of *high-touch, wide-tech, and professional depth*, and (b) a positioning strategy of high-quality *and* low-cost (i.e., exceptional *value*). The college slowly became known for providing a "private school education at a public school price."

Quality Milestones

By 1992, following numerous curriculum and faculty upgrades and a $5+ million renovation of Kepner Hall (its instructional facilities), the college's revised mission was paying significant dividends. The college reached its first major quality goal by earning accredited status from AACSB International—The Association to Advance Collegiate Schools of Business (AACSB), and became the first public university in Colorado to earn AACSB accreditation in both business administration *and* accounting.[15]

In 1999, in conjunction with a $10.5 million commitment from the Monfort family, the college was renamed the Kenneth W. Monfort College of Business (MCB). The gift was designed to provide a "margin of excellence" for programs at the college and honored a Greeley native and long-time supporter. Mr. Monfort was widely known as a pioneer whose commitment to innovation and quality through ethical business practice was legendary.

67

In 2000, the college earned Program of Excellence (POE) status from the Colorado Commission on Higher Education (CCHE).[16] The POE is a highly selective and prestigious award given only to programs that demonstrate widespread excellence and a readiness "to take the next step toward national prominence."

In 2004, Colorado Performance Excellence,[17] one of 40 Baldrige-based state quality programs in the US, recognized the college with its Timberline Award for performance excellence in approach and deployment. Later that year, MCB became the first college of business ever to earn the nation's highest award for performance excellence—the Malcolm Baldrige National Quality Award (MBNQA).[18] Established by President Ronald Reagan in 1987, the MBNQA program focuses on enhancing organizational effectiveness through the development and implementation of a comprehensive performance assessment system.[19] MCB's decision to pursue the MBQNA in 2002 was based on three key beliefs. First, it was believed that the formalization of the college's culture of continuous improvement would further strengthen program quality over both the short- and long-term. It was also believed that earning a state quality award, and ultimately the MBNQA, would help position the college as a quality leader. Finally, college leadership maintained that the BHAG (big hairy audacious goal) of earning the MBNQA would help motivate and align faculty and staff in the quest for continuous improvement over the coming decade.

The Journey Continues

Today, MCB is housed within UNC, a publicly-supported residential university of 12,078 students, offering a wide range of graduate and undergraduate degree programs in five academic colleges. Located on UNC's 236-acre campus in Greeley, Colorado (2000 Census pop, 76,930), MCB's primary service is offered to its 1,145 undergraduate majors. The college's educational services are delivered almost exclusively through a residential, on-campus learning mode of face-to-face student/professor contact. Class sizes (average of 30) are designed to enhance student/professor interaction. Distance educational delivery through technology is limited to the role of augmenting resident student classroom experience through use of ancillary techniques such as threaded discussions for extended class discussions, Web-recorded lectures for post-class reviews, and course-based Web sites with portals to related information sources.

Stakeholder Focus

MCB's mission and values focus singularly on pursuing excellence in *undergraduate-only* business education—a rare position among its regional and national peers. The college is one of just five undergraduate-only programs nationally to maintain AACSB accreditation in business and accounting. A leader in value when compared to its regional competitors, MCB's product

quality and learning environment is designed to exceed those of its peers. The *Denver Post* has described the college as "possibly the best bargain in business education anywhere in the U.S." In addition to its attractive price, MCB's commitment to a program strategy of *high-touch, wide-tech,* and *professional depth* has made it a value leader. This strategy is designed to meet the requirements and expectations of MCB's single key market segment (see Figure 2).

Figure 2
MCB's Key Market Segment—Business Majors

Key Requirements/Expectations

- Strong reputation of college and/or faculty
- Outstanding educational value
- Strong reputation of major/area of study
- Financial feasibility (affordability)
- Financial aid/scholarships (assistance)
- Outstanding facilities/technology
- Extra-curricular options (student clubs, speakers, conferences, and competitions)
- Course availability (scheduling)
- Outstanding placement for graduates
- Interaction with practitioners

High-Touch

Smaller class sizes are designed to facilitate faculty-student interaction in the classroom. No "mass sections" are permitted to ensure this interaction occurs across the entire curriculum. Smaller class sizes also allow for experiential, hands-on learning techniques to be employed and are designed to increase active learning levels within the student population. In addition, classes are taught primarily by doctorally-qualified faculty (83% of business core sections in Fall 2004), and no classes are taught by graduate students.

Wide-Tech

Since the Kepner renovation in 1987, MCB has invested millions of dollars in its technology infrastructure to support a curriculum that exposes students to a wide array of existing and emerging business technologies, enabling graduates to make a seamless transition into the workplace. The curriculum integrates technology within course content, and MCB prides itself on incorporating the most current versions of industry-standard technologies.

Professional Depth

MCB values professional business experience as a selection trait for its faculty. The college also utilizes an innovative E*xecutive Professor Program* to strengthen classroom currency and ties with the employment community

for graduates. Many of these professors are regionally- or nationally-known executives teaching in-residence, while others are brought to campus as visiting lecturers. The college also has developed partnerships with the business community to provide students with additional opportunities to gain real-world experiences through course components (e.g., business plans, advertising campaigns, market research, and portfolio management).

Exclusive attention to MCB's key market segment of students would ignore other important MCB stakeholders (see Figure 3). Each group has different, sometimes competing, but not necessarily mutually-exclusive, needs. A high-quality business program, both in reality and in reputation, forms the common thread for each stakeholder group.

Figure 3
MCB's Key Stakeholder Groups—
Primary Requirements/Differences

Alumni - Enhanced program reputation for adding value to MCB business degree

Employers - Access to well-prepared business graduates

Faculty & Staff - Fair compensation and opportunities for professional growth and development

Strategic Challenges

Driven by the college's mission and vision and guided by the stakeholder requirements described in Figures 2 and 3, MCB seeks recognition as the standard of excellence in Colorado for preparing business leaders and professionals. In so doing, MCB has prioritized four strategic challenges: overcoming a difficult fiscal environment; maintaining a high-quality faculty; maintaining high-quality student recruitment; and developing a market reputation to match existing program quality. To guide its progress in meeting these challenges, relevant performance indicators are identified and tracked. Where possible, these performance indicators are benchmarked against key competitors and/or national benchmarks.

Difficult Fiscal Environment

A recessionary economy, coupled with a state constitution that limits the rate of recovery through tax revenues, provided a 2003-04 budget allocation for UNC equivalent to that of fiscal year 1996-97. MCB absorbed a 10% overall budget cut for 2003-04, and prospects are bleak for rebuilding that lost revenue through state sources. Furthermore, economic woes have lessened the short-term ability of the private sector's potential for financial support.

70

MCB must now seek additional outside resources and increase its efficiencies while protecting program quality (e.g., reasonable class sizes, new technology). Relevant performance indicators include annual state and private funds available.

Maintaining a High-Quality Faculty

During the years, the college has successfully recruited and retained a high-quality, seasoned, and professionally experienced faculty. Maintaining quality in light of a national shortage of business PhDs and a senior faculty from which a number will retire over the next decade remains as a challenge (AACSB, 2002). The College must continue working to retain high-quality executive professors, recruit quality replacements when needed and will likely increase its proportion of executives in the future. It will also need to continue improving its faculty development and the evaluation system to guide faculty behavior into productive channels. Relevant performance indicators include student overall evaluations of faculty quality and student learning results, and faculty satisfaction and intellectual contributions.

High Quality Student Recruitment

Strategically, in order to recruit and retain a high-quality and diverse population of students, the college must continue to build reputation, while simultaneously working within Colorado's student-quality driven enrollment management system. Relevant performance indicators include quality of incoming student recruits and transfers, student retention rates, and overall student satisfaction.

Market Reputation Matching Program Quality

There is a strong connection between marketplace reputation and meeting the other strategic challenges. Increased funding from public and private sources, as well as faculty/student recruitment and retention, is tied to a stakeholder perception that program quality is high. MCB has identified and embraced this challenge by developing a new program of managing external relations through improved communications and partnerships. Relevant performance indicators include employer, parent, and alumni satisfaction, as well as quantity of media placements (in-process measure).

Student-Centered Process Framework

MCB's mission, vision, and values articulate the principles for its continuous improvement activities. The management system is built around a framework of systematic decision-making, organizational collaboration, and uniform commitment to mission that adheres to a student-centered process framework (see Figure 4). Beginning with student recruits and ending with graduates entering the workforce, this framework provides a tool for MCB leaders to focus their energies on those key areas that create stakeholder

71

value. Key performance indicators (KPIs) have been established to track overall organizational performance and to guide the College's daily operations. A second layer of indicators, SPIs (Secondary Performance Indicators), has also been constructed to provide a more detailed account of organizational performance. Figure 4 portrays how MCB's twenty KPIs map to the student-centered process framework. The college's 40 SPIs, though not pictured in Figure 4, similarly map into the framework. KPIs and SPIs measure achievement, satisfaction, and quality across each of MCB's key stakeholder groups. Each indicator has been tied to one- and five-year measurable goals that are reviewed annually to assess progress and opportunities for improvement, and measures are compared against national benchmarks where available and appropriate. In general, MCB sets a goal of being in the top 10% of any given comparator group.

KPIs, as identified in Figure 5, provide a mechanism for tracking overall MCB organizational performance. SPIs and other direct and indirect assessment measures provide a more detailed assessment of MCB student learning.

Figure 4
MCB Student-Centered Process Framework
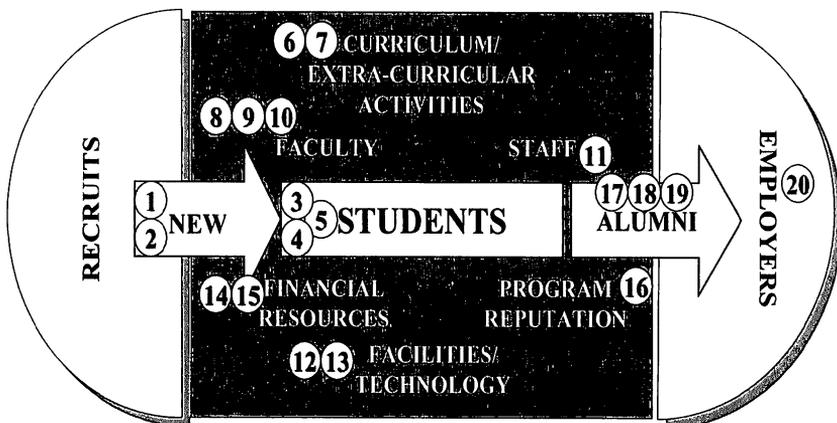
The tables and charts on the following pages provide examples of how MCB's student-centered process framework drives results, with data provided on a sample of the KPIs representing each strategic category listed in Figure 5. Referring to the framework's input (far left of Figure 4), student recruits are a critical component of the process. Though better students are not required to improve performance, better students can certainly enhance business programs in a number of ways. One would expect students with higher ACT scores, for example, to perform better on subsequent exit exams at a program's end. Though not formally studied as of yet, MCB's experiences suggest this

72

Table 5

Key Performance Indicators (KPIs) of Organizational Performance

KPI	Strategic Categories
1. Quality of incoming freshmen students (average ACT)	Recruits
2. Quality of transfer students (average transfer GPA)	Recruits
3. Student retention rates	Students
4. Business major counts	Students
5. MCB current student satisfaction (% recommending)	Students
6. Student learning in business (average overall ETS MFT score)	Curriculum
7. High-touch curriculum (average class size)	Curriculum
8. Quality of faculty (% academic and/or professional qualification)	Faculty
9. Quality of academic faculty (graduating student satisfaction with quality of instruction and faculty)	Faculty
10. Faculty program satisfaction (average)	Faculty
11. Staff satisfaction (average overall)	Staff
12. Student satisfaction with facilities/computing resources (average)	Facilities/technology
13. Faculty satisfaction with computing resources (average)	Facilities/technology
14. Total available state funds (annual)	Financial resources
15. Total available private funds (annual)	Financial resources
16. MCB press coverage (media coverage generated)	Program reputation
17. Placement of graduates (% employed full-time)	Grads/alums
18. Graduating student satisfaction (average)	Grads/alums
19. Alumni satisfaction (average)	Grads/alums
20. Employer satisfaction (proportion satisfied)	Employers
Source: Monfort College of Business 2004 Strategic Plan	

correlation to be far less than a perfect 1.0. Side-effects of selectivity, nonetheless, include enhanced prestige for the program and pride by those who successfully enroll and complete these programs. Moreover, most instructors will agree that increasing the quality of students in the classroom enhances the learning environment for all students. In 1999, MCB introduced a Freshman Finley Fellow program to attract higher quality student recruits. The program offers incoming scholarships to talented freshmen as a means

of gaining an initial commitment to MCB. Successful freshmen are then well-positioned for a number of other scholarships, of which there are far more opportunities available for continuing students. Figure 6 offers evidence of success over time in attracting a higher-quality student population.

Figure 6
MCB Student Quality—Proportion of Entering Freshmen ≥24 ACT

Finley Program Impact Begins

1998 1999 2000 2001 2002 2003 2004
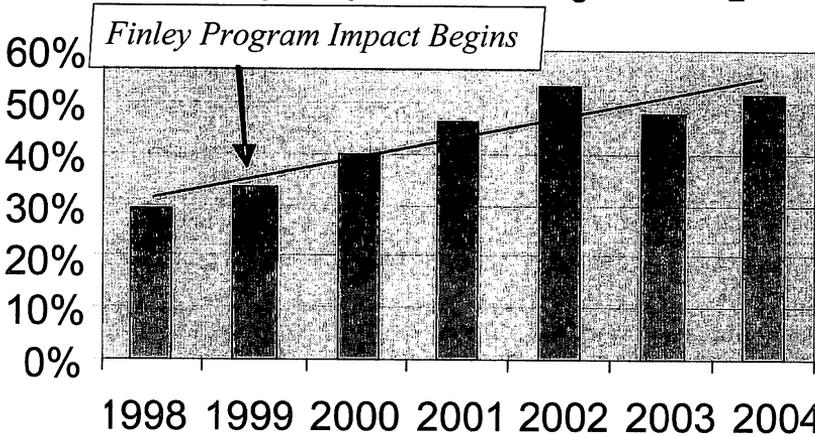
Source: University of Northern Colorado, Institutional Research and Planning

While each of the previous measures is important, none is particularly meaningful if student learning is not occurring. Of the 450 total points available (on a 1,000 point scale) in the Malcolm Baldrige Educational Criteria for Performance Excellence "results," one third are specifically tied to student learning. Learning outcomes are also central to the AACSB standards and simply cannot be over-emphasized (AACSB, 2005). A myriad of direct and indirect measures are available to business programs—all with their own distinct advantages and drawbacks. MCB currently uses both direct and indirect measures of student learning. The Educational Testing Service, Inc. (ETS) provides a nationally-normed examination of core business skills that is currently used at nearly 500 business schools (ETS, 2004).[20] The national benchmarks provided by this exam are useful, both for program management, as well as for communication of program quality to external audiences.

Figures 7 illustrate two measures of MCB's relative overall performance on the ETS exam over a five-year window. In-process measures for the 2004-05 academic year (Summer and Fall 2004) show continued improvement for MCB graduates and represent a cumulative increase of nearly 30 percentile points since 1993-94. Furthermore, the percentage of MCB graduates performing in the top 5% and 10% bands is nearly triple the national averages. To date in 2004-05, over 82% of MCB graduating seniors have scored above the national mean. Figure 8 demonstrates how MCB has increased student performance on the ETS exam in every content area within the five-year

Figure 7
MCB ETS Student Performance

6-Year Proportion of MCB Students in Top Percentiles
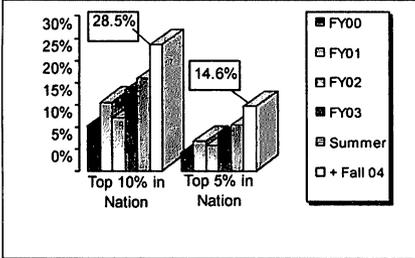
ETS Overall Performance
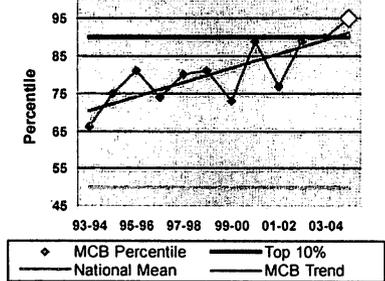
window. For 2003-04, MCB students performed in the top 20% in every content area. In-process measures for 2004-05 place MCB students in the top 10% in every area except economics (top 15%). Such results resonate very strongly with prospective employers, as well as with student recruits and their parents and high school counselors.

Stakeholder satisfaction, though not a direct performance measure, is generally linked to student learning outcomes. MCB student satisfaction with the undergraduate business program has historically been quite high on a national scale. Overall student satisfaction has continued to place MCB in the upper 1% of business schools participating in the EBI survey. Contributing

Figure 8
MCB ETS Area Performance—All Students Combined
for Each Area (Percentile)

Area	99-00	00-01	01-02	02-03	03-04	5 year +/-
Acct.	68	83	70	85	95	+27 pts.
Finance	78	86	74	88	90	+22 pts.
Mgmt.	81	91	78	80	85	+4 pts.
Mkt.	87	97	97	88	90	+3 pts.
QBA	78	89	80	90	90	+12 pts.
Interntnl.	48	56	65	79	80	+32 pts.
Leg./Soc.	69	57	58	76	85	+16 pts.
Econ.	41	65	56	79	80	+39 pts.
MCB.	73	89	77	89	90	+17 pts.
U.S.	50	50	50	50	50	-----

to the satisfaction level is MCB's *wide-tech* strategy which requires that close attention is paid, with managed support processes, to ensuring wide availability of state-of-the-art software and hardware. Survey data indicate strong faculty and student satisfaction with regard to hardware and software support, as well as with the availability of MCB technology.

Employers are aligned with the output side of MCB's student-centered process framework. With the assumption that satisfied employers will want to continue recruiting MCB graduates, high placement rates and evidence of excellent learning outcomes are important measures of success. To address a gap in information with this important stakeholder group, MCB developed its own employer survey. First-year results indicate that employers rated MCB more favorably than its comparative peer set by a margin of 3-33 percentage points.

Comparable regional and national undergraduate placement rates, particularly for individual majors and business schools, can be difficult to acquire. In many cases, the data may be available, but the response rates, or the timing of the surveys, make exact data comparison impossible. As a result, MCB uses two benchmarks, overall UNC placement (excluding MCB) and the June Colorado unemployment rate to measure the success of its students in securing employment. MCB has consistently outperformed both benchmarks.

Using Assessment Data to Further MCB's Strategy

As described earlier, MCB's decision to focus on an undergraduate-only mission has been central to the development and implementation of a high-quality program strategy. This focus has allowed for a concentration of financial and personnel resources at one program level. However, choosing an undergraduate-only focus also has limited MCB's ability for adopting many of the reputation-building strategies followed by its regional and national peers. For example, the majority of ranking programs (e.g., *Wall Street Journal*) are geared exclusively to MBA programs. Even the *U.S. News & World Report's* undergraduate business program rankings illustrate the difficulty of smaller and/or regional schools rising to the top quintile ranks. The rankings, based on the perceptions of business deans around the US, are arguably influenced greatly by the much more visible MBA and PhD programs of many schools.

As a result, after almost two decades of building program quality, MCB reached a point where its performance results had placed it in rather elite company on an "anonymous" basis. In other words, while the college continued to generate top quality results, it remained a relative unknown even in its own region. In 2002, MCB adopted its fourth strategic objective—developing a "market reputation that matches its program quality."

Because reputation is important to stakeholders throughout MCB's student-centered process framework, MCB increased its emphasis on public relations by hiring a Director of External Relations prior to fiscal year 2003.[21]

The newly-created position is held by an executive with significant, national public relations experience. MCB established this position by combining part-time instructional and executive professor funds—consistent with MCB's high-touch strategy, this individual normally teaches a full load. The college witnessed an immediate boost in press coverage the following year, and in-process measures for 2004-05 have exceeded the previous year's totals in only the sixth month.[22] News articles highlighting MCB student performance on the ETS exams, MCB student wins in national competitions, and notable research by MCB professors have begun to enhance both the regional and national reputation. Clearly, the selection of MCB as a 2004 MBNQA recipient has been central to the increase in press coverage in the current year and will offer further opportunities for the college to share its key messages and results with a variety of stakeholder groups not otherwise possible. It is also helping MCB more credibly build a market reputation that matches its program quality. Testimonials from highly-placed "experts" are obviously helpful as well.

Lessons Learned and Advice from the Monfort College of Business Journey

Over MCB's 20-year quality journey, senior leaders have learned a myriad of important lessons. Perhaps the best way to summarize these lessons is via a "top 7" list:

1. Establish a mission-driven focus - A mission-driven entity, business school or just business, is simply more likely to succeed in the long run if it pursues an market-supported, focused strategy. The decision by MCB faculty and senior leadership in 1984 to eliminate the college's master's-level and doctoral-level programs was dramatic, difficult, and contrary to prevailing trends. Nevertheless, the regional competition and the opportunities for building excellence at the undergraduate level supported the decision. The ability of the college to *stick with* this focused strategy for over twenty years through three deans has clearly been a primary reason for MCB's current level of success. A less dramatic, but important part of this strategy was to reduce college offerings to a single undergraduate degree. By designing a curriculum *focused* first on fundamental business skills and second on specialization, the college's curriculum has enjoyed greater longevity than many programs.

2. Close the loop - A consequence of the college's achievement of AACSB accreditation in 1992 was a voracious appetite for the collection of data. Unfortunately, like many organizations in education and business, MCB was not consistent in its review of all data and in acting on the information—what is often referred to as "*closing the loop!*" MCB's adoption of the Baldrige criteria, among other management decisions, has been central to the college's improvements in recent years. The management-by-fact

philosophy has focused senior leaders' attention on a smaller number of key performance indicators and has greatly enhanced their ability to make necessary process changes in response to changes in the data.

Recent examples of "loop closing" at MCB include (a) a curriculum change (an additional required course) in one of our majors after ETS data indicated that students in this major did not perform well in finance; (b) a change in the content of one of our technology courses to replace existing instruction on a software package (MS Front Page) with one that was more highly rated by employers and our alumni (Macromedia Dreamweaver); and (c) based on EBI survey feedback that students were not satisfied with the availability of technology and support, we built a new student lab and hired a technology director for our College.

3. Adopt a quality mentality - Economic woes are a common thread for most institutions of higher education over the last decade. Colorado's unique constitutional structure has magnified the impact of tighter state budgets to higher education in general. Unfortunately, many taxpayers fail to see the importance of higher education, with many viewing such expenditures as discretionary. It is the authors' assertion that, over time, higher education can make a positive impact in changing these perceptions by promoting the establishment of and accountability to measures of our learning outcomes and organizational effectiveness. The MBNQA and the *results-oriented* assessment it is geared toward is one system that can provide higher education with a highly respected alternative to combating some of these negative public perceptions.

4. Integrate the Baldrige system - Similar to an accreditation process, the MBNQA application process requires an applicant to conduct a rigorous self-study of not only the processes used to drive their organization, but of the results of its activities over an extended period of time. Past recipients, and many who have not won the award, claim that they apply, not with the goal of the award in mind, but with the clear objective of receiving the *feedback* of the well-respected Baldrige examiners (i.e., volunteer quality experts from a variety of industries across the US). For educational applicants, a nominal application fee results in a minimum of 200 hours of review, and site-visited applicants receive over 1,000 hours of review from a team of 8-12 experts.[23] The bottom line is that for any readers who have made it this far in the chapter, it is likely that said individuals are seeking to improve their organization. If truly serious about accelerating improvement, the Baldrige criteria and the overall MBNQA system will aid those efforts considerably.

5. Incentives matter, but support the faculty - At the core of any successful business program is a strong, committed faculty. AACSB-accredited business colleges need to consider their mission as well as AACSB standards when defining faculty expectations for both scholarship and teaching. Moreover, rewards should be available for faculty exceeding minimum expectations. For example, utilizing externally-raised funds, MCB established

professional-development grants that faculty can earn for publications and for receiving various college and university awards. These grants allow faculty to travel, to purchase software or data, or to support other teaching or research-related activities. Because all of these activities *enhance faculty skills* and are aligned with MCB's mission, the grants simultaneously enhance faculty's human capital and improve the business program.

6. Getting central-administration buy-in Business school deans expend considerable time and energy in pursuit of private funds to support current and future programs. Nevertheless, most public and private schools receive their core funding through a centrally-administered funding process. This common model necessitates the following approaches when building strong partnerships with central administration. First, MCB established *initial buy-in* to the process by educating central administration about the Baldrige process. The college invited a representative from a previous MBNQA winner to campus to meet with faculty and central administration. The purpose of this visit was to get buy-in both "upstream" and "downstream." Following the visit, MCB was able to get the approval of the provost to substitute MCB's application to the MBNQA for the normal university program review. Next, the college *solidified the buy-in* by sharing not only the substantive feedback with faculty and central administration, but, perhaps more importantly, results of improvements were also shared directly with those stakeholders and were indirectly echoed through significantly increased media exposure. The media exposure not only helped central administration see the value of the Baldrige process, it helped further MCB's progress toward one of its strategic objectives.

7. Keep it up. MCB, you won the Baldrige Award, are you going to Disney World? - In the midst of celebrations over the Baldrige Award, we were asked the equivalent of that question. One might think the temptation to relax would be strong after receiving the MBNQA. Even if MCB's quality mindset were insufficient, however, AACSB's new assurance of learning standards would very quickly eliminate any such temptation.[24] The new standards not only require that each program create specific learning goals, but further require that programs measure learning as it relates to these goals and define improvement cycles when learning outcomes are unsatisfactory. By having a single program at the undergraduate level, MCB has positioned itself favorably relative to most schools, which offer multiple business degree programs. Currently used direct measures (e.g., ETS, MFT) and indirect measures (e.g., EBI surveys)[25] of learning outcomes provide solid evidence of MCB's effectiveness. Nevertheless, for AACSB purposes and for continuous program improvement, MCB will need *to identify new direct measures of learning outcomes.* Devising these new measures in a way that minimizes resource requirements (or securing new resources to support the new measures and associated processes) is likely to provide the most significant challenge to business colleges as they seek to gain initial or maintain AACSB accreditation in the upcoming years.

Endnotes

15 For more information on AACSB, including a list of accredited schools, see http://www.aacsb.edu. Last accessed January 17, 2005.

16 For more information on CCHE, visit their Web site at: http://www.state.co.us/ cche_dir/hecche.html. Last accessed January 21, 2005

17 See http://www.coloradoexcellence.org/ for more information on Colorado Performance Excellence. Last accessed January 17, 2005.

18 See http://www.nist.gov/public_affairs/releases/2004baldrigewinners.htm (Last accessed January 17, 2005) for a general winner's announcement, or http://www.mcb.unco.edu/web/gi/Baldrige/ (Last accessed January 17, 2005) for more specific information on the award announcement.

19 See http://www.quality.nist.gov (Last accessed January 17, 2005) for more information on the Malcolm Baldrige National Quality Program or Calderon, Green, & Hackness for a description of the Baldrige process in education.

20 See the ETS Web site at http://www.ets.org/hea/mft/compare.html for further information. Last accessed January 17, 2005.

21 Fiscal years begin July 1 and end June 30.

22 To view some of the media coverage generated visit http://www.mcb.unco.edu/ web/gi/nr/pressClippings.cfm. Last accessed January 21, 2005.

23 See "Why Apply" at http://www.quality.nist.gov/Why_Apply.htm for more information on the value of the Baldrige system. Last accessed January 17, 2005.

24 For more information on the new AACSB accreditation standards see http:// www.aacsb.edu/accreditation/standards.asp. Last accessed January 17, 2005.

25 Though EBI surveys provide indirect measures of learning outcomes, some argue that such measures of self-perceived learning are highly accurate (85-93%) indicators of actual learning. See Wortham & Harper (2003) at http:// www.webebi.com/Research/ENG/eNews/FeaturedArticle/PDF%20Folder/ BLcorrect.pdf for a discussion of this point and further references.

References

AACSB International. (2002). *Management education at risk: Report of the management education task force to the AACSB – International Board of Directors*. Retrieved August 02 from http://www.aacsb.edu/publications/ metf/default.asp.

AACSB International. (2005). *Eligibility procedures and standards for business accreditation*. AACSB International, Tampa, FL. Retrieved Jan 05 from http://www.aacsb.edu/accreditation/standards.asp.

Bush, G. W. (2004). Message from the president to recipients of the 2004 Malcolm Baldrige National Quality Award at http://www.nist.gov/ public_affairs/releases/Baldrige04_PresidentStatement.htm. Last accessed January 19, 2005.

Calderon, T. G., Green B. P. & Hackness, M.. (2004). *Best practices in accounting program assessment*. Teaching & Curriculum Section, American Accounting Association, Sarasota, FL.

Colorado Performance Excellence. Denver, CO. Retrieved January 17, 2005 from http://www.coloradoexcellence.org.

Educational Benchmarking Incorporated. Retrieved January 17, 2005 from http://www.webebi.com/ManagementEducation/Index.htm.

Educational Testing Service. (2004). *Major field test in business.* Retrieved January 17, 2005 from http://ftp.ets.org/pub/corp/hea/kbusdata.pdf

Jones, D. (2004, November 26). Malcolm Baldrige Winners Met and Set Goals. *USA Today*, McLean, VA.

National Institute of Standards and Techonlogy (NIST), Malcolm Baldrige National Quality Program, Washington D.C. Retrieved January 17, 2005, from http://www.quality.nist.gov.

National Institute of Standards and Techonlogy (NIST), Malcolm Baldrige National Quality Program, Washington D.C. *Why Apply?* Retrieved January 17, 2005, from http://www.quality.nist.gov/PDF_files/Why_Apply.pdf.

Monfort College of Business, University of Northern Colorado, Greeley, CO. *Malcolm Baldrige National Quality Award for Education 2004 Application. Retrieved January 19, 2005, from* http://www.mcb.unco.edu/web/gi/Baldrige/baldrige2004public.pdf.

Monfort College of Business, University of Northern Colorado, Greeley, CO. *Faculty/Staff Handbook*. Retrieved January 17, 2005, from http://www.mcb.unco.edu/web/fs/FSHB.pdf.

Worthman, K. & Harper, V. (2003). Learning outcomes assessment. Retrieved January 17, 2005, from http://www.webebi.com/Research/ENG/eNews/FeaturedArticle/PDF%20Folder/BLcorrect.pdf.

Authors Bio

Timothy E. Jares is Associate Dean and Associate Professor of Finance in the Monfort College of Business at the University of Northern Colorado. Combining an industry background in systems engineering with academic training and research in investments and corporate finance, Jares brings a unique perspective to assessment and problem solving in a business school environment.

Joe F. Alexander is Dean of the Monfort College of Business at the University of Northern Colorado. A long-time advocate for fact-based management, Alexander has earned a reputation for fostering organizational quality through consensus-building leadership and strong collaboration with key internal and external partners

The AACSB AOL standards require that program learning goals include both knowledge and skills. At first, it may seem that assessing students' knowledge would present relatively few problems; after all, traditional classroom assessment methods have been doing this for centuries. As the author points out below, however, this is certainly not the case. Among the issues facing business schools when choosing between assessment methodologies is whether to "make or buy." Standardized tests for business knowledge are readily available (at a price), but do they offer advantages over "home grown" instruments? The discussion below provided a side by side comparison of the widely used ETS field test with locally-developed examinations to help schools match the appropriate method with their goals, resources and capabilities.

CHAPTER 6
ASSESSING BUSINESS KNOWLEDGE
Denise M. Rotondo
Salisbury University
Perdue School of Business

Background, Mission, and Goals

At the end of the day, every business program must face the undeniable fact that its core purpose is to produce graduates who have a foundation of business knowledge. Each student graduating from a school or college of business, regardless of major or academic concentration, should carry forth a broad understanding of business concepts, theories, processes, and relationships that distinguishes a business student from other undergraduate majors. It would seem that a goal so fundamental to the academy should be among the simplest to assess for program improvement. Yet, this is certainly not the case.

Perhaps the difficulties stem from unclear property rights, because the content knowledge set of business belongs to everyone and no one among a business faculty. Perhaps faculty members have less of a psychological investment in other related skill areas, and they scrutinize those assessment measures with far less rigor than measures of business knowledge. Or perhaps perceived threats to academic freedom underlie the tension. No matter the reasons, decisions about how to assess core business knowledge and how to use assessment data for continuous improvement in a business program are among the most controversial and strongly contested decisions made in the name of assurance of learning.

In this chapter, I present an overview of the process for assessing core business knowledge. The merits of common assessment measures will be explained. Using examples from several programs, successful approaches to assessing business knowledge are illustrated and compared. I give an extensive discussion of the key issues related to implementation and closing the loop. Finally, I address some of the pitfalls to avoid so that the process of assessing business knowledge is as productive as possible.

Learning Goals

One distinguishing feature of business education is the emphasis placed on knowledge and skills imbued within students' academic majors juxtaposed with the knowledge and skills across all business majors. A reflection of the integrated and multi-functional nature of business, business students are exposed to a broad array of foundation knowledge and functional knowledge in addition to the depth of knowledge and development of skills within each major. In a school of liberal arts, students majoring in history must meet the academic demands of the history program. In most cases this doesn't include an expectation that students will demonstrate knowledge and comprehension of literature, art, or philosophy outside what may be included in general education requirements. In business programs, however, finance majors must know something about marketing, management, production operations, accounting, information technology, and the like, because at least some knowledge of those disciplines is necessary to successful careers in finance.

Across the board, most business programs have at least one learning goal expressing the expectation that graduates will acquire business knowledge across the range of disciplines. Many schools extend beyond understanding to the level of application of knowledge to new problems or situations in a business context. Learning goals involving general business knowledge normally target the most basic levels of learning (e.g., under Bloom's taxonomy, knowledge, comprehension, and possibly application). This is simply because most business students only receive a general or "principles-level" exposure to business disciplines outside their major. The following learning goals capture the essence of business knowledge in a variety of programs:

1. Upon graduation, a student will demonstrate knowledge of core business concepts and topics, and be able to apply that knowledge to new problems and situations (Salisbury University).
2. Students will know, apply, and integrate the content in one's major and will apply and integrate accumulated cross-discipline concepts (Towson University).
3. Students will know the core concepts within each business discipline, marketing, accounting, finance, management, and information technology (Appalachian State University).

Once a learning goal is written, it becomes tempting to jump directly into a debate about what particular assessment measure should be chosen. It is assumed the learning objective would read something like:

Students will take a comprehensive test of business knowledge and score an average of 75%.

This is because the method most commonly used for assessing content knowledge is a test, and the only question presumed to remain is, "What performance level meets expectations?"

Moving straight to decisions about a measure would be premature at this point. It results in one of the most prevalent problems with the assessment of business knowledge today. That is, identifying an instrument you wish to use and then trying to find a way to overcome its shortcomings, or worse, fitting your curriculum to the measure itself.

It is necessary to be certain precisely (a) <u>what</u> business concepts or topics should be taught and learned before deciding and (b) <u>how</u> you can ascertain if indeed the learning has occurred. To say a student will know "marketing concepts" is too broad. The faculty must define the scope and content of the business knowledge to be learned and assessed in each discipline for two reasons. First, faculty own the curriculum. Faculty are charged with delivering courses and creating learning opportunities for students in line with an existing mission and business curriculum. Whatever assessment method is ultimately selected must measure the concepts deemed essential by the faculty teaching within the program, not the concepts deemed appropriate by others.

Second, the assessment data gathered, regardless of the method, must be useful for program improvement. Gathering data that is of marginal value to the faculty teaching in a particular program is unproductive at best. Even less desirable is being pressured to make curricular changes in order for your students to perform better on a test that was anointed for assessment before determining what would be assessed.

The first step in defining the essential business concepts is to query the faculty. Many schools have an institutional history of carefully reviewing their curriculum to ensure proper coverage and consistency across multiple sections of core courses. Other schools have a culture that places such an emphasis on academic freedom that the concept migrates into freedom to teach personal preferences, resulting in idiosyncratic learning experiences independent of the stated curriculum. If a thorough review of the curriculum and core concepts has not been done recently, begin the process by asking faculty (by department of discipline) to create a list with no more than 5-7 key concepts (including first-level and second-level sub-headings). The list need not represent every concept that a student should know at the completion of a course, but rather what most students (and in particular majors from other

business disciplines) should know and retain until graduation. Asking for faculty participation and securing their endorsement of what constitutes "essential business topics and concepts" is critical to avoiding conflict downstream once assessment data have been gathered.

After defining what is to be assessed, it is necessary to begin the tasks of defining measurable objectives and identifying the method and/or means by which students' achievement in business courses will be measured. (I am making an assumption, here, that the core curriculum indeed covers what will be assessed.) The following section reviews alternative methods to directly assess business knowledge.

Measurement Options: Test A vs. Test B

Although there are a variety of methods that can directly assess students' knowledge of business, standardized tests or "common exams" are most prevalent. Faculty are familiar and comfortable with using tests to assess student learning in their courses and may be skeptical that any other method could be valid. Nevertheless, you must select the best method to assess business knowledge within the context of your program's mission and values, for your curriculum and your students.

Methods and Options: Student portfolios, simulations, and performance activities (within a capstone class or an assessment center) can serve as direct measures and capture data on student knowledge of business concepts, as can standardized tests. The selection decision should balance the set of knowledge and concepts previously identified by the faculty against financial resources, technological resources, time constraints, and against the intended uses of the data.

As stated earlier, tests of business knowledge are the most common assessment method across the board and, thus, will be the focus of discussion in this paper. The tests are normally comprised of multiple-choice questions, tend to be relatively easy to administer and score, and the results are easy to interpret without subjectivity. Tests of business knowledge yield copious quantitative data, useful for all sorts of analyses of student performance and program effectiveness.

What tests of business knowledge lack is an ability to provide a more subjective or qualitative assessment of student performance. Further, they provide little or no opportunity to assess higher-level thinking (e.g., analysis, synthesis, and integration). Since most learning goals related to core business knowledge do not require students to demonstrate higher-level thinking, the latter shortcoming is far less relevant. The issue over the usefulness of quantitative, objective data as compared to subjective or qualitative insight may be very relevant though. An analogous comparison most educators would appreciate is the difference in the depth of learning one can infer from student responses to essay questions versus that evidenced by the correct selection among four alternatives on a multiple-choice exam. Open-ended

items allow for a more subjective analysis of what a respondent has learned where objective items sometimes capture what a respondent recognizes or recalls when prompted. In the interest of space and because there is a clear preference for objective tests of business knowledge, I will focus on the issues of test choice, implementation, and use of results.

The question of the hour seems to be whether to use standardized, commercial tests or to create a standardized "home-grown" test. Table 1 provides a quick summary of the advantages and disadvantages of adopting the Major Field Test of Business offered by Educational Testing Service compared to developing a test exclusively for your program. The ETS Major Field Test of Business (from this point forward referred to as the Field Test) is the most popular commercial test available, used by over 460 different institutions. It contains 120 multiple-choice questions measuring knowledge of accounting, economics, finance, management, marketing, quantitative business analysis, and law as well as social and international concerns. Locally developed tests are less prevalent, but do exist to meet the assessment needs of single programs and university systems alike. Both options (a commercial test or a locally developed test) have advantages and disadvantages, and it important to give careful consideration to all factors before making a decision. The ETS Web site (http://www.ets.org) effectively explains the details of the Field Test, its proper uses, and the like. Thus, in this chapter I will focus upon those issues most central to the assessment of core business knowledge and continuous improvement.

Advantages of Using the ETS Major Field Test: Without question, the greatest advantage of the Field Test is that it allows a school to compare the performance of its students to the performance of students in other business programs who took the test that year.[26] Scores for the Field Test are scaled, normed, and reported in percentiles, which enables an institution to benchmark performance.

Because the Field Test is widely used (469 institutions for academic year 2003-2004), the marketing potential of being able to report that "…our students scored better than 80% of all business students in a national test of business knowledge" is significant.

The ability to have a national benchmark against other business schools was a key reason why the University of Northern Colorado[27] chose the ETS Field Test of Business. Being the first winner of the Baldrige National Quality Award in education, the national comparative data and the ability to track improvements in student performance over time was important to UNCO's success. Dr. Tim Jares, Assistant Dean, said that the business school began by setting two goals relating to the Field Test. First, no discipline would be below the 60th percentile in overall student performance, and second, no discipline would have more than two years of a downward trend in student performance on the Field Test. After a series of curricular changes, student performance improved, and the school's new goal is to be at or above the 90th percentile overall.

Table 1
Comparison of the ETS Field Test of Business
to a Locally-Developed Test

	ETS Field Test of Business	Locally-Developed Test of Business
Advantages	• Quick and easy to implement • Provides comparative results that can be used to benchmark institutional performance against other business programs • Affords institution the ability to leverage student performance for marketing or promotion of program • Used by many business programs, both accredited and non-accredited • Assesses eight business areas	• Provides results that assess core knowledge that corresponds directly to the program's curriculum • Greater flexibility for analysis of results • More conducive to using results for curricular improvements • Opportunity for test to evolve as curriculum evolves • Can be designed to fit within existing class schedules
Disadvantages	• Cost per student of exam plus additional interpretive reports is high, can be prohibitive • Requires two-hour time block to complete, typically administered outside of class time • Often difficult to use data from results to evaluate curriculum, especially when institutional performance is at or above the mean • Norm-based exam with percentile results are often misunderstood; as norm shifts each year, so do percentiles • Mean student performance for correct responses is low in most disciplines, reflecting low student motivation and/or substantial degradation of knowledge • Test questions change infrequently and may not align well with a single institution's curriculum, making continuous improvement difficult	• Unknown external validity, therefore no comparative statements can be made regarding student performance • Requires careful attention in development in order to establish internal validity • Test development takes faculty time, therefore not cost-free, but can be lower than Field Test • Requires careful handling of data to maintain integrity of results and allow for accurate analyses

87

National benchmarks can be particularly useful to smaller programs, according to Dean William Rupp at Montevallo University. Explaining why his school uses the Field Test, he emphasized that high student achievement on the Field Test allows a small school to demonstrate the quality of its curriculum and its faculty. When it can be said that "our students did as well or better than..." students at larger and/or more well-known business programs, the ability to attract students to your institution and program is boosted tremendously.

There is no question that the Field Test has an advantage over locally developed tests of business knowledge in external validity, useful for promoting a business program through comparative statements about student achievement. Its 120 questions have been established as both reliable and valid indicators of student achievement in the areas tested. However, the AACSB standards for assessment and assurance of learning do not require schools to benchmark student achievement.

Many proponents of the Field Test as an assessment tool for business knowledge cite the "reliability and validity" of the test to argue against a test developed locally. Debates about the psychometric properties of local versus commercial tests, though, are often a bit misplaced when the objective is to assess student knowledge in your program for AACSB accreditation purposes. There is a qualitative difference between internal validity and external validity. Without the external validity inherent in the Field Test, generalizing results of student learning in your business program to what students in other programs learn is next to impossible. If marketing and promoting the relative strength of your program is of strategic importance, the Field Test may be the best method for assessing core business knowledge. If the marketing and promotion potential is not critical, a locally-developed exam can be just as effective for assessment purposes.

Disadvantages of the Field Test of Business: For each student who takes the Field Test the institution must pay $26 (as of the date this chapter was written). And, the various reports that can help a school make better sense of student results represent an additional cost. For larger programs, the expense of the Field Test can be significant. For smaller programs, the Field Test may simply be cost prohibitive.

Another noteworthy disadvantage is that Field Test scores are scaled, normed and reported as percentiles. Without going into the messy details, a simple example can illustrate the problems this can cause. Assume that you give the Field Test to your students in the upcoming spring term and learn that they have scored in the 60th percentile, overall, and your weakest area was finance where your students scored in the 45th percentile. These results are based on the performance of your students and all other students who took the Field Test in that Spring as well. Seeing a "problem" in finance, you express your disappointment to the finance faculty and charge them with reviewing the core class in financial management to improve student learning.

Next spring, a year later, you administer the Field Test again and find that your students have moved to the 65th percentile, overall, but are still at the 45th percentile in finance. To the uninitiated, it would appear your students have done better (except in finance) which is great news. Or is it? First of all, the number of participating institutions against which your students are being compared may have gone up or down, so it's difficult to say whether your students are really doing better or whether other students have done worse. As for finance, (which happens to be the discipline where students everywhere perform the worst), it is entirely possible that your students increased their percentage of correct answers in the finance area. But since almost every school is working to improve student learning in finance, your students may have simply kept pace as the percentiles go. The important fact that students gave more correct answers in finance is what should really be celebrated, but it is more likely someone will wonder if the finance faculty really made any changes at all. The mean percentage of correct answers for finance questions on the Field Test is around 37%[28] (meaning students getting 37% correct are in the 50th percentile for finance). And, to score in the 95th percentile in finance, the student only needed to answer 49%[29] of the finance questions correct. It is challenging to pinpoint why changes in student performance on the Field Test occurred, but multiple factors are at play.

It may be surprising to note that many people, even academics, have great difficulty interpreting percentiles. Percentiles are moving targets, and the numbers have nothing to do with how many questions (or concepts) students got correct. Being in the 95th percentile doesn't mean a student got 95% of the questions correct...the student could have answered as few as 56% of the accounting questions correctly to be in the top percentile in 2003-2004. A student who answered 56% of the questions correctly on an accounting test would get an F in most courses...so being in the 95th percentile may not be much to brag about.

Consider as well the following comments shared on the AACSB assessment listserv regarding student performance and percentile ranks: "...for the economics area in Spring 2003, a school whose students averaged 8 correct on the 20 economics questions (40%) received a percentile rank of 25, while a school with an average of 9 correct was in the 55th percentile, and one with an average of 10 correct (50%) was in the 80th percentile." That students get such a low percentage of questions correct (except in management and basic statistics) suggests that either their motivation to do well is very low, that students have retained very little of the basic knowledge we expect of them at graduation, or that the questions are entirely too difficult (unlikely, but a possibility).

Putting aside the interpretation of scores, there have been a number of recent changes made at ETS to address other deficiencies in the Field Test for AACSB assessment purposes. One example is a new report made available in Spring 2004 that provides an analysis of the percentage correct for each

test question (there is an additional cost for securing this report). The analysis can be used to "fine tune" curricular improvements. Without this information, a school only knows how well its students scored in each discipline (e.g., in the 55th percentile for marketing). Without this information, the marketing faculty must guess at what topics or concepts, in particular, the students missed on the Field Test. How should they go about improving the curriculum? Should they cover the 4-Ps more extensively? Or should the role of consumer perceptions be emphasized more? Having item-by-item information allows the faculty to see if there are weaknesses in a particular area or just across the board. Because the information must be requested separately and involves additional cost, it is considered a disadvantage of the Field Test.

Advantages of Locally Developed Tests: Most programs that have developed a local test have done so for one or more of the following three reasons (a) because they could not afford to use the Field Test (it was too expensive); (b) because the test questions and/or content did not match the business curriculum students were taught; or (c) the logistics of arranging a two-hour block of testing time for students was too difficult.

Towson University in Maryland developed a test of business knowledge for its program primarily because the business faculty felt a locally developed test could assess what was taught in their curriculum more effectively and yield results more helpful to curriculum improvements. The cost of the Field Test was a factor, but secondary to concerns about the appropriateness of the content. A committee of faculty experts with full support and encouragement of the dean developed the test. The committee worked with the various departments to draft and refine questions for an exam that is now given as part of the capstone course. Although student performance on the business test cannot be compared with results at other business schools, the questions on the test itself are reliable items, and students' scores can be considered valid indicators of their learning achievements in Towson's business program.

The test is used to improve the program, but such improvements are only possible with some effort. Towson University has a plan to communicate the results to the dean and the college's Assessment Committee. The results will be broken down by course and according to student major, so that student learning for each core course can be compared as a whole and by academic major. Ideally, the results will motivate faculty and departments to improve their teaching, thereby increasing student learning. The major problem to be faced is avoiding singling out specific faculty members or groups of faculty for retribution. If this were to happen, the faculty buy-in to the assessment process will probably lessen. Assessment experts are adamant that data collected to evaluate student learning in a program should never be used to evaluate faculty.[30]

The Towson model is a good example of why concerns about test reliability and validity are sometimes misplaced. Because the intent at Towson

90

was *not* to use student scores for national benchmarking or to market the program, the external validity of their local test is a moot point. The process Towson used built a reliable and internally valid exam, so the school can make effective use of students' scores to help improve the business curriculum.

The California State University System is an example of taking a locally developed test to the next level. The Cal State System created its own test of business knowledge that became available to all system schools in academic year 03-04. The decision to create a state-wide test was made by the group of California's AACSB business school deans. The development process involved faculty across the state writing, reviewing, and refining questions to establish a reliable test. Cal State Long Beach provides the administrative support for the test, including scoring and reporting of results to the various campuses. Associate Dean Mo Kahn at Cal State Long Beach explained that the two-hour time block needed for the Field Test simply didn't work within their existing class-time structures and precluded many working students from being able to attend the test administration. The Cal State System test was designed to be completed in 70 minutes (and can be administered in two 35-minute parts). Dr. Kahn also said the high cost of the Field Test was another reason why a locally developed test was needed. The cost to a campus using the Cal State test is $2 per student (compared to $26 per student for the Field Test). Included in the cost are summary and comparative reports, so schools can analyze results between campuses. Dr. Kahn stressed the advantage, too, of being able to provide reports that are completely customized to the needs of a particular campus.

An inherent advantage of a locally developed test is that a school can generate extensive and customized reports on the results. One can know, if so inclined, the response frequency for each answer option for every question on the test. It is also possible to perform additional analyses not possible with the Field Test to see if performance is affected by demographic or academic variables (such as the number of hours transferred into your institution).

Disadvantages of Locally Developed Tests: Faculty time and maintenance of the data top the list of disadvantages for developing a local test. Faculty must be involved in the process of writing, editing, refining, piloting, and revising test questions within each area to be tested. And faculty time is not free. It is important to have a dedicated champion in the initial stages and a clearly defined process for administering the test and handling the data.

Although you can collect a number of demographic and academic data in order to provide more depth to the analysis of results, you will need to handle and maintain the data somewhere. Who will have access to the data, what process will be used to capture it, and where it will be housed are questions that must be answered. In addition, the means by which the data are secured and reported to avoid inappropriate uses of results must be

91

established. Should these responsibilities fall on a person or committee, the additional duties must be appropriately factored into work load.

Making the Choice: What to do? A good starting point is to compare the topics included on the Field Test with the list of essential business concepts your faculty consider important to the curriculum. Securing such a list was suggested earlier as part of the establishment of learning goals and learning objectives. If the list of essential concepts identified for your program doesn't match well with the Field Test, the Field Test may not be a good choice for your assessment method. You will be testing students on material different from that taught in the program and it will make improving the curriculum extremely difficult.

If the Field Test is used, an excellent way to make the test more useful for assessment is to add your own questions. Schools have the opportunity to include up to 50 additional questions on the Field Test. Doing so allows a program to assess areas not covered by the Field Test, to test certain area in greater depth, or to include more applied questions. It is my personal opinion that any institution using the Field Test should take advantage of this option.

Regardless of whether a school uses the Field Test or develops a local test, there are a number of implementation issues worth considering because of their potential impact on the usability of the results. The next section addresses the key concerns.

Implementation Issues

What to do? A good starting point is to compare the topics included on the Field Test with the list of essential business concepts your faculty consider important to the curriculum. Securing such a list was suggested earlier as part of the establishment of learning goals and learning objectives. If the list of essential concepts identified for your program doesn't match well with the Field Test, the Field Test may not be a good choice for your assessment method. You will be testing students on material different from that taught in the program and it will make improving the curriculum extremely difficult.

If the Field Test is used, an excellent way to make the test more useful for assessment is to add your own questions. Schools have the opportunity to include up to 50 additional questions on the Field Test. Doing so allows a program to assess areas not covered by the Field Test, to test certain area in greater depth, or to include more applied questions. It is my personal opinion that any institution using the Field Test should take advantage of this option.

Regardless of whether a school uses the Field Test or develops a local test, there are a number of implementation issues worth considering because of their potential impact on the usability of the results. The next section addresses the key concerns.

Student Motivation

Administering a test of business knowledge can be done as a stand-

alone activity outside of a particular course. Many programs, though, incorporate the Field Test or their own test of business knowledge into the program capstone course. Schools have invented myriad ways of factoring a student's score on the business knowledge test into the final grade in the capstone course. A number of programs don't incorporate a student's score into any academic grade. Regardless of the particulars, one of the paramount concerns with the implementation of these tests is the impact of student motivation on the accuracy of the test results. Without results you use, you cannot complete the ultimate task of closing the assessment loop.

Will My Score Count? Most people who have taught in higher education have learned that if an assignment or test doesn't affect a student's grade, there's almost no chance the student will complete the work or prepare for the test with the intention to do well. The riskiest approach to assessing business knowledge with a test is to appeal to a student's sense of duty to the program or school, personal pride, or sense of accomplishment. Few students are motivated by these intrinsic attributes of success; such an approach would yield dubious results at best. If the integrity or validity of the test results are questionable, they should not be used for curriculum improvements and "closing the loop." Thus, a student's performance on the assessment test should probably "count" in some way toward academic grades. How much it should count is another matter.

Even if the score on a test of business knowledge is a component of a course grade (such as a capstone course), students make rational decisions about the amount of effort to supply on the test. Students will put more study time and preparation into a test worth 15% of a grade than will be allocated for a test worth 1% of a final grade. As stated earlier, there is no consistency across business programs whatsoever in how much test scores affect student grades, or whether different faculty teaching multiple sections of a single course treat test scores similarly. This inconsistency often stems from claims of academic freedom or different philosophies about the validity and usefulness of comprehensive business knowledge tests. Regardless of the cause, it is highly possible that students in Professor A's class will be strongly motivated to do well while their friends in Professor B's class will not. Therefore, for the overall good of the program, faculty should be encouraged to reach consensus on how the students' scores will be calculated into final grades.

Complicating the problem of motivation to perform well is the fact that these tests cover material that students may not have seen in over two years. For example, the Field Test is comprised of a fairly significant percentage of questions on economics and accounting. Many business students are exposed to these disciplines through lower-division courses taken as a sophomore (sometimes a freshman) and take the Field Test as a graduating senior. Without continual reinforcement and repetition in the curriculum, memories fade quickly to say the least. It shouldn't be surprising to find students performing poorly on those sections. For these reasons, a student's

score on any test of business knowledge should not determine if the student can graduate from the program. Nevertheless, it is important that students do their best on the test so that a host of unneeded curriculum changes aren't put in place. Dealing with fading memories becomes an implementation challenge in search of a solution.

When to Give the Test? The time lapse between when core courses are taken and when the business knowledge test is administered could potentially be solved by giving the test earlier in the student's academic program. The University of Northern Colorado is considering a change that would require students to take the Field Test just after the completion of the junior year rather than in the senior year. While the concepts tested would be somewhat fresher, one could argue that our concern should be what the students retain upon graduation. Moving the test administration back from the final semester will likely raise questions about why a final assessment is needed at all, and that invariably leads to a series of discussions about why passing grades in the core courses shouldn't suffice for assessment purposes. To satisfy the spirit and intention of assessment within the standards for assurance of learning, measures of business knowledge should be taken at or near the end of a student's business program.

Coaching the Players? Another controversial matter is whether students should be prepped for the test. The answer depends on what you seek to understand about your students, your business program and your curriculum. In many schools, faculty hold review sessions to help clear out the cobwebs and refresh memories prior to taking the test. Others argue this is "gaming the scores," and artificially inflating what students really know about business concepts. So, if you seek to capture the "residual knowledge retained," students should take the test without a refresher course. If the intent is to see what students in your program are capable of achieving, offering review sessions would make sense. A "compromise" solution may be to offer passive review support through Web-based tutorials or other written materials, requiring students to show a bit more self-reliance and initiative to access and use the information.

There is no empirical research to date that links student performance on the Field Test (or any business knowledge test) to one or more motivation strategies. Insufficient motivation to do well on the Field Test is a potential explanation for the low mean correct scores. Table 2 summarizes the major conclusions for addressing student motivation. Anecdotal evidence also suggests administrators and faculty can encourage better performance through managing perceptions and rewarding high achievement. To do so, the test must be given its proper place in the program.

The next sections address strategies useful for boosting student motivation controlled by the administration and faculty.

Administrative Support

Part of the Culture: One common element found in schools that have used the ETS Field Test with success is a strong emphasis placed on the importance of test performance in the school's culture. At Northern Colorado, Dr. Tim Jares highlighted the fact that the dean and administration were visible, vocal supporters of the test among the student body. Students at UNCO clearly understood that their scores on the Field Test mattered—not just for the school, but also for them as individuals. High individual scores can be used by students to signal their potential to prospective employers. And, many programs issue new releases or include stories in school publications about student performance on the ETS Field Test, touting high overall achievement and the achievements of high scoring individual students. Building a strong culture around the importance of doing well on the Field Test is key to influencing student perceptions and, hopefully, test results.

Students seeking business degrees at Montevallo University are inculcated with the importance of doing one's best on the Field Test of Business through several strategies. First, Dean Rupp holds a breakfast for students the day the test is administered where he speaks to them about how important their test scores are for the college. The student who scores the highest is issued the "Dean's Scholar Award" which is presented at a faculty meeting by Dean Rupp. Although there is no financial component to the award, the recipient receives a certificate and has his or her picture taken with the dean.

Karen Constantino, from ETS, shared a list of ideas for motivating students gathered from the various institutions who use the Field Test. The various strategies seek to build enthusiasm and reward performance. Students who perceive that the test is important and believe that the faculty and administration care about the scores are likely to try harder to do well on the Field Test. Among her suggestions—sending letters to freshmen or sophomores letting them know about the test they will be taking as a senior. This emphasizes the importance of retaining business knowledge as students progress in the program. The use of rewards (both large and small) also encourages more effort. Among the various rewards mentioned by Ms. Constantino are gift certificates, food, raffles for gifts or cash, a month of parking privileges, and receptions with university administrators or employers. Also noted are the various coercive methods used to ensure all students take the test, such as registration holds and failed graduation audits for those students who don't show up.

Faculty Buy-In

The faculty represent the front line when it comes to student perceptions and motivation to perform. If faculty support tests assessing business knowledge through their words and deeds, the effect on students will likely be positive. If, on the other hand, faculty are determined distracters, students

will almost certainly under-perform. Among the implementation issues, a key concern is whether faculty should review the test questions or not.

Many hold the strong belief that if faculty review the test questions, they will teach to the test and taint the results. Dean William Rupp at Montevallo University encountered a debate about faculty review of the Field Test that became so divisive, one's position on the issue became an indicator of moral or personal character rather than professional philosophy.

Educational Testing Service makes review copies of the Field Test available to faculty, and goes even further in support of the position that faculty should review the contents of the test. ETS encourages any school considering the adoption of the Field Test for assessment purposes to carefully review the contents and determine if there is a sufficient match between the test contents and the curriculum of the program. The purpose in using the test for assessment is not helped or purified by "surprising" students or testing them against material they have not been exposed to. Indeed, faculty should be confident that learning opportunities are provided for the vast majority, if not all, of the topics included on the Field Test.

Should a program decide to develop a test locally, at least some faculty members will most definitely be involved in determining its contents. As long as the intended use is for curriculum improvement, this is not a flaw or a breech of professional ethics. It is part of the content validation process.

If there is any real issue of concern relating to faculty review of the test contents, it would be the threat those actions have on the legitimacy of external comparative statements. That some programs "teach the test," "prep their students," and then celebrate the performance of their students can be a real concern for schools that adopt a hands-off approach when assessing their students. Neither approach is wrong…they simply serve different objectives. However, it begs the question at schools electing to forgo the added miles to facilitate top performance, "Is our performance really lower, or did other schools perform higher?" Since external benchmarking of student performance is not required for accreditation maintenance, such questions are tangential. The primary concern among the administration and faculty should be how the results from the test can be used to make curriculum improvements. Table 2 summarizes the basic suggestions for implementation.

Closing the loop is the holy grail of assessment activities. The most pristine assessment plan is worthless unless it has been used to improve the business curriculum. The process and issues related to using knowledge tests for continuous improvement are discussed next.

Closing the Loop

No matter what method a school selects to assess business knowledge, the data gathered must be used to help improve the business curriculum, enrich students' learning experiences, and raise the quality of the business

Table 2

Obtaining Useful Results: Getting Students, Administrators, and Faculty to March in Concert

Implementation Strategies for Improving Student Motivation
• Students' scores on tests (including the Field Test) should count for a meaningful portion of a grade (or should have the potential to impact their academic performance) • Faculty should be consistent in how scores are factored into grades across multiple sections • The test (or assessment of business knowledge) should be administered at or near the end of a student's business program • The more active the faculty in helping students re-learn concepts prior to an assessment, the more the results reflect the maximum achievement possible within the current curriculum
The Role of the Administration
• Simply stated, the importance of "doing one's best" on assessment tests must become part of the institution's culture • The dean, associate dean, assessment coordinators and department chairs should be visible and vocal supporters of assessment • Providing student rewards and recognizing high achievers underscores the importance of taking the test seriously
The Support of Faculty
• Faculty are the face of assessment to students; if students perceive that the faculty think assessment tests are a waste of time, students won't put forth effort to do well • Faculty absolutely should review the topics and associated questions included on the Field Test; ETS strongly encourages this as essential to the validation process

program. Quite literally for business knowledge tests, the results must feed back to the faculty who deliver the curriculum. If arguments over academic freedom didn't occur earlier, they will surely surface when the test results point toward the need for curriculum change that's too close to home. Faculty can be very supportive of the need for change as long as it's another department that needs to change!

If the need for change is determined through the use of a normed knowledge test developed by outsiders, faculty seeking to resist continuous

improvement efforts can draw upon a list of logical arguments to support their case.

- The Field Test really doesn't fully capture what we feel must be taught in our department.
- The areas on the Field Test where students are weak aren't that important in our courses.
- Many changes in the field have taken place, and the Field Test just hasn't caught up yet.

These and similar arguments can be minimized, if not eliminated, by engaging in the proper due diligence during the planning phase of assessment. Faculty must identify and agree upon what is to be assessed before selecting a measure to use for assessment. The convenience and ease of implementation provided by the Field Test can be intoxicating to programs seeking to move assessment forward quickly. Unless the tradeoffs have been given careful consideration up front, problems are simply being shifted downstream to be addressed later when the cost of starting over is much higher.

Concerns about closing the loop aside for a moment, a number of schools have made changes to their curricula-based upon results from student performance on the Field Test of Business. The approaches taken by those schools are varied and worth examining. Ohio Northern University is familiar to many because it is highlighted on the AACSB Web site. The University of Northern Colorado has a similar story. Both schools targeted areas where their students scored relatively poorly on the Field Test for curriculum revision. Ohio Northern, for example, raised its students' quantitative scores from below median to the 91st percentile by increasing the emphasis and exposure to math and statistics concepts in the business curriculum.

Northern Colorado has an even more pointed illustration of how curriculum changes improve learning. The Field Test results showed that both their marketing and management students were testing much lower than desired in the area of finance. The management department changed the curriculum of the major and required a second finance course. The marketing department chose not to do so. While the scores on the finance questions improved significantly for management students (as measured by the Field Test), the marketing students' scores have not.

Dr. Tim Jares, at Northern Colorado, added that one challenge he sees looming in the future for his school is continuing the improvement efforts once the students' reach the 95th percentile in an area. At both UNCO and Ohio Northern, the areas of the curriculum needing initial improvement were relatively easy to identify because students' scores were low. When this is the case, small changes in the curriculum of core courses may translate into dramatic improvements on the Field Test. Because the Field Test's upper

limit is the 95th percentile, curriculum improvements become much more difficult to make after the "low hanging fruit" has been picked. The only solution is to delve into greater detail by drilling down to the individual topic level. And, as closing the loop gets closer to changes at the individual topic level, the potential for faculty resistance and fights over academic freedom becomes greater.

Another factor that presents challenges in making curriculum improvements—regardless of what test is being used—is the plight of transfer students. Micro- and macro-economic concepts are typically taught in the lower-division, and many students have taken these courses at a community college and not at the baccalaureate institution. The same is true for courses addressing financial and managerial accounting principles. It becomes particularly challenging to address curricular weaknesses in these areas unless upper-division economics and accounting courses are required or unless students receive tutorial assistance. Suggesting to the accounting faculty that, based on test results, they need to address the weaknesses in managerial accounting concepts will require those evaluating continuous improvement efforts to adjust expectations for future outcomes.

If the Field Test is the method being used to assess business knowledge, additional caution is needed to avoid over-interpreting the results. The difficulties of percentile scores were discussed earlier, so it is important to look for other "evidence" a particular curricular change is in order before the change is implemented. Dean William Rupp summarized this well by stating that, "The Field Test is not a stand alone. It must be part of an assessment program."

Further, when making changes in an effort to improve test scores, avoid instituting a set of changes simultaneously. If a school sets out to improve student motivation on a test while also making changes to the curriculum, it will be impossible to disentangle the effects each action contributed to the next set of results.

Curriculum improvements without specific details of what concepts or topics are weak should not be expected to yield dramatic improvements in a short time frame. If faculty have not reviewed the Field Test and if the institution does not obtain the item response summary report available through ETS, the goal for faculty asked to improve student learning becomes, "Just do your best" or "See if you can do any better." That is hardly the kind of motivation to which faculty respond. Similarly, observed changes in student scores do not necessarily mean the faculty have successfully addressed the deficiency. Trends over time or triangulation of multiple data would be more robust indicators.

Lastly, be sure that curricular improvements are indeed improvements. An example will illustrate the point. As a professor of organizational behavior, I find the questions measuring management knowledge on the ETS Field Test to be quite simplistic. (That is entirely my opinion, and I acknowledge

that my concern may be equally likely if a locally-developed test were used at my school that contained questions similar to the Field Test). If a review of student scores on the Field Test at my school revealed that more coverage of management history was needed in the principles class, I must admit I would question whether that constitutes an *improvement* of the curriculum. The ability to apply management concepts and theories is much more important to me than accurate recall of definitions. As a professor, I believe I would be doing my students a disservice by diverting class time away from application toward memorization of terms. But, as a steward of the school's continuous improvement efforts, I am doing a disservice to the school by obstructing the changes if our students are assessed by the Field Test. Herein lays the dilemma of academic freedom as perceived by faculty.[31]

The accreditation standards ask us to engage our students in meaningful learning experiences while making improvements to the curriculum in our business program. An assessment test of business knowledge should be one of several measures used to find out just what it is your students know about core business concepts. Borrowing the wise words of so many before me, it cannot be stressed enough that while waiting for the perfect assessment test to come along, you will miss out on opportunities to benefit from good assessment tests.

Pitfalls to Avoid

Resistance to curriculum improvements becomes intense if assessment data have been carelessly or inappropriately interpreted, which happens more frequently than one might think. Consider the following story.

At a school I know, the Field Test was administered for three years as a result on one strong advocate among the faculty willing to administer the test and report the results. As it happens, the faculty advocate for the Field Test was notorious for being critical of faculty in the departments of marketing, economics and finance. After each round of tests results came in, he quickly tabulated student performance by subject crossed by student major. Overall, the school's students scored between the 82nd and 87th percentile. But invariably, there were one or two majors performing below the 50th percentile in marketing, economics, and/or finance. At every opportunity possible, be it faculty meetings or hallway talk, this Field Test advocate would mention what a poor job those economics and finance faculty were doing or how the marketing class must be a joke. Even when challenged about the interpretation of percentile scores, when reminded that almost 50% of the school's students took their economics courses at community colleges, this fellow was unrelenting. When it was pointed out that students were doing equally worse on legal issues, he quickly changed the subject. The victims were quite content to simply ignore the misplaced attacks until the new standards were approved and the ETS advocate joined the assessment committee.

Still steadfast in his views, he has now been quieted by the other faculty

who quickly learned the finer points of interpreting percentiles and who carefully read the assurance of learning standards. There is a now a full-court press for a locally developed test and an active resistance against using any existing Field Test data for continuous improvement action plans.

The faculty advocate for the Field Test committed a cardinal sin of using assessment data and in particular of using the Field Test data. He used the results to "evaluate faculty performance" rather than student learning. Such sins turn faculty against even the best assessment method and against assessment in general. Assessment data from tests of business knowledge can not be used as a way to grind axes or reveal flaws in the faculty.

The example also underscores the importance of data security and data access. An established procedure for gathering, analyzing, and reporting student assessment data should be in place before the first test is administered. This helps to avoid information falling into the wrong hands. Had such a procedure been in place at this school, the results would not have been open to the interpretation of one person.

Conclusion

Assessing business knowledge requires travel on a straightforward path with numerous potholes that must be negotiated. Many programs have adopted the ETS Field Test of Business as their primary method of measuring knowledge of core business concepts. Other schools have developed their own test. As long as a school's faculty are comfortable with the decision, either choice can yield the results can be useful in pointing out potential weak spots in a curriculum. Neither, though, should be the single measure of business knowledge. For those schools that develop their own tests of business knowledge, don't be swayed by those who criticize local tests for their "lack of validity." There are enough criticisms that can be leveled at the Field Test that I can only call the contest a "draw." As of now, there is no perfect test of business knowledge. What matters most is how the data gathered from the measure are used to improve the curriculum.

In this chapter, I have reviewed the basic process for assessing business knowledge. Once a learning goal has been established and a learning objective defined, the difficultly of selecting, implementing and utilizing a measure of business knowledge begins. Whereas many programs see this as perhaps the "easiest" learning goal to tackle for assurance of learning purposes, I can state with confidence it is not. Because the "business content" goal lies so near and dear to the core of what faculty do in a business program, the devilish details to close the loop often take us by surprise. I encourage all programs to anticipate problems in advance and give careful consideration to the issues expounded in this chapter. As well, I anticipate a future when non-test alternative methods of assessing business knowledge are better developed and more widely accepted in the academy. Until then, get the faculty to support a "good" method and put it to work for your program.

Endnotes

26 The use to individual students in evaluating their own performance, while important, is not addressed in this chapter. Rather, the focus here is on how a program could use data reported for its institution and for all institutions for program assessment.

27 Editor's Footnote: The role that assessment played in UNCO's successful application for the Award is discussed in the chapter by Jares & Alexander (Vol 1: No. 1).

28 Based on the 2003-2004 data reported by ETS for the Major Field Test of Business.

29 Based on the 2003-2004 data reported by ETS for the Major Field Test of Business.

30 There are two problems associated with using program assessment data to evaluate individual faculty members: First, learning related to program goals takes place across the curriculum; thus, it is unfair to penalize a single faculty member for a poor result. Second, there is a conflict of interest in asking a faculty member to gather data on student learning that will then be used to evaluate him or her.

31 Editor's note: For a discussion of academic freedom, and its implications for assessment of student learning, see H. James William's chapter in this volume.

References

Personal interview with Tim Jares, University of Northern Colorado, 12/21/04.

Personal interview with Mo Kahn, California State University Long Beach, 1/9/05.

Personal interview with William Rupp, Montevallo University, 1/8/05.

Correspondence from Karen Constantino, ETS, National Accounts and Special Projects

www.aacsb.edu/resource_centers/assessment/tools.asp
www.aacsb.edu/resource_centers/assessment/printf.asp
www.ets.org

Author Bio

Dr. Rotondo is the Associate Dean for the Franklin P. Perdue School of Business at Salisbury University in Salisbury, Maryland. The Perdue School of Business was recently reaccredited under the new AACSB International standards, where as Associate Dean, Dr. Rotondo planned and initiated the school's assessment and assurance of learning process. Dr. Rotondo is also an Associate Professor in the Department of Management, where she teaches courses in Organizational Behavior and Human Resource Management.

Many business schools are recognizing the critical importance of developing students' managerial and interpersonal skills. Yet, it is not enough to attempt to teach the skills. There must also be efforts to measure skill development. Assessing student interpersonal skills serves two primary functions— student development and program assessment. In this chapter we clarify what is meant by managerial interpersonal skills and suggest possible assessment approaches including assessment centers and 360-degree feedback.

CHAPTER 7
ASSESSING THE UNASSESABLE: INTERPERSONAL AND MANAGERIAL SKiLLS

William H. Bommer, Cleveland State University
Robert S. Rubin, DePaul University
Lynn K. Bartels, Southern Illinois University-Edwardsville

Background, Mission, and Goals

In the past decade, an increasing focus on skill development in undergraduate and graduate business curricula has emerged (e.g., Bigelow, 1995; Bigelow, Seltzer, van Buskirk, Hall, Schor, Garcia, & Leleman, 1999; Boyatzis, Stubbs, & Taylor, 2002). A number of important factors are likely to have contributed to this heightened awareness of skill education. The now well-known research by Porter and McKibbon (1988) clearly illuminated the necessity for business schools to improve their ability to teach management, leadership and other interpersonal skills. Similarly, surveys of corporate recruiters routinely cite interpersonal and leadership skills at the top of the skills most desired in graduates (e.g., Eberhardt, McGee, & Moser, 1997). Recently, well-known management scholars have been highly critical of the overall value of most MBA programs (Pfeffer & Fong, 2002) and cite their relative lack of practical and economic merit. Further, the American to Advance Collegiate Schools of Business (AACSB) publishes accreditation standards that encourage schools to support their claims that skills are being inculcated with formal assessments measuring those skills. Though these standards used to be quite broad, the most recent AACSB standards require institutions to complete assurance of learning measures that determine *direct* educational achievement (Thompson, 2004). Thus, institutions which state that they train certain managerial skills must present primary evidence that those skills have indeed been learned.

In response to these trends, many schools now teach managerial skills and offer skill-based management courses (Bigelow *et al.*, 1999) or, at the very least, supplement their traditional organizational behavior and human resource management courses with skills activities and exercises. If

management skills are to be developed, a key component of that process involves feedback. Students need opportunities to practice their managerial skills in realistic situations and receive feedback on those skills (Whetten & Cameron, 1995). Skill assessment is important in identifying current levels of competence and serves as an important catalyst for change. As students practice their developing skills, it is important to provide ongoing feedback. Feedback should be based on objective, accurate and credible measurement of skills. Thus, assessment of managerial skills plays a key role in management development.

Although many institutions are beginning to integrate skill-based education, rigorous skills assessment reflective of the skills being taught has been slower to develop (McConnell & Seybolt, 1991; Riggio, Mayes, & Schleicher, 2003) and for good reasons. Rigorous evaluative techniques are not necessarily "intuitive," thus making it difficult for academics and administrators to distinguish best practices from malpractices. In addition, much confusion exists regarding what constitutes managerial skills. For example, many confuse student personality or attitudes with managerial skills. This confusion regarding what constitutes skills often leads to mistakes in assessing those skills.

Further, despite overwhelming evidence that deployment of management skills via human resource management practices has a strong impact on organizational productivity and performance (e.g., Huselid, 1995), the practices remain characterized as "soft," "elusive," and "unassessable," and are often seen as without practical merit (Rynes & Trank, 1999). Thus, resources granted to "soft skill" evaluation as part of a larger program of student assessment may be viewed as unnecessary or unattainable. Rynes Trank, Lawson and Ilies (2003) remarked, "…we are a long way from showing that students who take courses in organizational behavior…behave any differently or perform more effectively than those who haven't" (p. 279). The lack of behavioral change evidence is partly a function of measurement issues. That is, it may be a testament to the difficulties that accompany assessment and development of interpersonal skills. Indeed, these skills are not only difficult to measure—they may take a lifetime to master.

Given the complexity present in developing management skills and the contextual factors presented above, the purpose of this chapter is two-fold. First, we seek to clarify what is meant by managerial skills and skill assessment. In particular, we will draw distinctions between the various types of assessment that exist and the purposes they serve. We devote most of our discussion to interpersonally-oriented managerial skills, as they are often seen as the most elusive and "unassessable." Second, relying on research and our practical experience in assessment, we elaborate on two methods for assessing interpersonal skills that have great promise for increased usage in academic settings: assessment centers and 360-degree assessments. These two methods have their roots in organizational assessment techniques

and show enormous potential to contribute to the educational assessment arena.

Defining & Assessing Interpersonal Skills

Virtually all business schools claim to develop some form of students' interpersonally related skills and rightfully so—interpersonal skills matter. Research by the Center for Creative Leadership (CCL) showed that among junior executives identified as "high potentials," many failed to be promoted into executive roles (i.e., were derailed) despite enormous resources provided to them. CCL found among these derailed high-potentials, there were great problems with interpersonal relationships, inabilities to build and lead a team, and failure to develop or adapt to change. These findings are not unique to the case of derailed executives. Research in the leadership literature, for example, has shown that specific interpersonally related leader behaviors are strongly related to employee performance and commitment (Judge & Piccolo, 2004). Further, among college students, Waldman and Korbar (2003) found that students with increased levels of interpersonal skills were more likely to have higher starting salaries, more promotions over a five-year period and higher job satisfaction. Clearly, interpersonal skills are important for students' future success, and resources spent towards development and assessments of interpersonal skills are fully justifiable.

The term "skill" is used to characterize a specific behavior or set of behaviors that can be learned and repeated with consistent results. Business students develop many skills throughout their studies, including highly technical ones such as financial calculations, economic analyses and software operation. These are learned through practice in the classroom and reinforced over time to provide consistent results. The constellation of skills known as "interpersonal" is no different; they involve specific behaviors performed with consistent results. The difference of course between technical skills and interpersonal ones is that interpersonal skills necessarily involve other people, whether through interaction or in the context of other people (i.e., managing one's self in relationship to others). Some behaviors that demonstrate interpersonal skills include clarifying the task, communicating clearly, and seeking input from others.

Interpersonal skills differ greatly from student knowledge and ability. Knowledge represents an organized body of information obtained through education or experience. Abilities represent aptitudes or capacities to perform a certain behaviors. Thus, skills represent some form of "doing." While it is true that knowledge and ability are sometimes prerequisites for learning and performing a skill, they are by no means direct measures of the skill itself. As Mintzberg (1975) noted some 30 years ago: "Management schools will begin the serious training of managers when skill training takes its place next to cognitive learning. Cognitive learning is detached and informational, like reading a book or listening to a lecture. No doubt important cognitive material must

be assimilated by the manager-to-be. But cognitive learning no more makes a manager than it does a swimmer" (p. 60).

Just as knowledge and ability do not represent direct measures of skills, neither do individual differences such as attitudes, values, motivations and personality characteristics demonstrate skill. These broad sets of individual differences represent a complex interaction between an individual's experiences in the world and genetic predispositions. Although some individual differences such as personality are often similar in focus to a specific skill, and may be a necessary condition, they do not constitute possession of a skill. For example, the personality trait of extraversion describes an individual's predisposition towards sociability and talkativeness. Extraversion may be quite helpful to an individual who is performing the skills associated with oral communication. However, extraversion itself does not guarantee possession of oral communication skills, nor ensure that an individual will execute the skills appropriately. The same can be said for attitudes and values. Possession of a certain attitude or value may facilitate the performance of the skill, but it is not a measure of the skill.

These distinctions above are not simply "splitting hairs"—they are important in that possessing an ability to perform a skill, being predisposed towards the skill, and/or knowing about the skill, are far different than actually *performing* the skill. When these distinctions are not carefully considered, the interpretation of skills assessment is likely to lead to flawed conclusions. That is not to say that personality measures, for example, are not useful. To the contrary, such measures hold enormous potential for helping students develop a strong sense of self-awareness about their preferences and behavioral tendencies, and we would recommend them as part of an overall development program.

Towards Interpersonal Skill Assessment

Business schools have excelled at assessing students' ability and knowledge while doing relatively little in the skills area. For example, we screen applicants closely, using ability measures such as SATs, ACTs, or GMAT scores. Similarly, we are well-equipped to test student knowledge through examinations of facts, what Anderson (1985) called *declarative knowledge*. Although critically important within a program of assessment, declarative knowledge measures focus heavily on cognitive-type learning (Bartels, Bommer & Rubin, 2000; Waters, Adler, Hartwick, & Poupart, 1983) to the exclusion of affective and skill-based learning (Kraiger, Ford & Salas, 1993). Here again, we are not suggesting that evaluation of students' "technical skills" (e.g., database management, financial calculations, etc.) is not important; rather, they are more readily measurable through traditional assessments and represent a great deal of declarative knowledge. Thus, it is imperative to closely examine what assessment techniques are tapping. For example, Table 1 lists many of the measurement methods used by business

schools to purportedly measure interpersonal skills. Several of these approaches measure behaviors directly (e.g., assessment center, 360-degree feedback, peer evaluations of class projects), while others clearly measure something other than behavior (e.g., personality inventories).

In thinking about the approaches to measuring interpersonal skills, we recognize that for many institutions, interpersonal skill assessment represents a tradeoff between administrative burdens and behavioral assessment. When comparing the different approaches to measuring interpersonal skills, it is important to consider the costs of administration. These costs may include time to develop or obtain, coordinate and administer the assessment process. Many of the paper and pencil techniques are readily available and can be administered and scored rather quickly. For example, personality inventories and peer evaluations of group projects may be easy to obtain or develop and easy to administer. The development of an interview is a straightforward task, but conducting individual interviews can require significant amounts of time; thus we categorize it as having moderate administrative costs. Other forms of assessment (360-degree feedback and assessment centers), on the other hand, can often be quite burdensome, requiring significant amounts of time to develop and administer. Thus we categorize them as being high on administrative costs.

In order to provide high quality, credible feedback it is important that the feedback is based on objective data. Self-report techniques (e.g., personality inventories, interviews and other survey techniques) may be biased by impression management, social desirability, ego-protection or other distortions. Self-evaluations also tend to suffer from high levels of leniency and low levels of agreement with ratings from other sources (Harris & Schaubroeck, 1988). Of course not all self-reports are designed equally, and some make considerable efforts to make the assessment as behavioral and objective as possible (Riggio & Riggio, 2001). Interviews by virtue of their human interaction provide a forum for assessing interpersonal skills unlike the paper and pencil approaches. In Table 1, we classify peer evaluations of class projects and 360-degree feedback as moderately objective approaches. These two approaches are improvements over self-ratings, but they involve ratings by individuals who may have limited ability or motivation to provide accurate and constructive feedback. In assessment centers, feedback is based on assessor observations of standardized work simulations. Assessors have no previous knowledge of the students and can rate the behaviors as performed in the exercises. Thus, two approaches that may be the most beneficial in terms of providing high quality assessment of management behaviors are the assessment center and 360-degree feedback. Unfortunately, these two approaches are the most administratively demanding. The remainder of this paper will present experiences with these two approaches used for assessing student managerial skills and suggest ways to minimize administrative costs.

Table 1
Comparison of Methods of Assessing Interpersonal Skills

Measurement Technique	Administrative Cost	Relative Objectivity	What is Being Measured?
Personality Inventories (e.g., NEO-FFI, MBTI)	Low (easily obtained and administered)	Low (Self-report)	Personality, temperament, preference, (not interpersonal skills)
Interviews	Moderate (easy to develop, time-consuming to administer)	Low to Moderate (self-report, easily faked, but the interviewer has the opportunity to view interpersonal skills firsthand)	Varied content may include past behavior, future intentions, communication skills, skill at relating to the interviewer
Peer evaluations of class projects	Low (easy to develop and administer)	Moderate (classmates may have limited ability to provide constructive and unbiased feedback)	Peers' perceptions of behavior and contribution toward group goal
360-degree feedback	High (substantial data collection and analysis and feedback effort)	Moderate (work colleagues who may have varied motives for providing accurate feedback)	Multi-source measurement of behaviors in unstandardized situations
Assessment Center	High (substantial effort to develop, administer and provide feedback)	High (unbiased assessors who are trained to provide constructive feedback)	Assessors' perceptions of behavior in standardized and realistic work simulations

Emerging Assessment Methods

In our pursuit of innovative methods to assess interpersonal skills, we have developed a few decision criteria that have helped us to determine which skills to assess and to evaluate skill assessment methods. First, if a skill can be evaluated well using some form of traditional means, we will not revisit it. For example, knowledge tests measuring applied or procedural knowledge (e.g., determining the appropriate accounting rule to apply) are generally quite good at capturing technical skills. Second, we have narrowed our assessment of interpersonal skills to those that are inherently observable and can be reliably measured through some form of observation. Thus, we seek to assess observable behavior. Third, for an assessment method to be truly useful, it must serve as a meaningful platform for providing feedback to

students about their development, while at the same time be useful to administrators/faculty for purposes of aggregation and program evaluation. In the end, we tend to err more on the side of student-focused feedback, which likely reflects our professorial role, rather than a bias towards program evaluation.

Applying these criteria, we have had great success in both promoting skill development and assessing student outcomes through the use of assessment centers (AC) and 360-degree assessments. In particular, we have been using the AC methodology for 8 years and have assessed over 20,000 undergraduate and graduate business students from all types of universities. Recently, we also began to use 360-degree assessments in order to capture certain skills not assessable through the AC. Below; we introduce the reader to both methods and provide detailed examples from universities employing them.

Innovative Assessment Methods
Assessment Centers

An assessment center (AC), despite its name, is an evaluation and development technique, not a location. The AC was used first by the U.S. military to aid in officer selection in World War II. Soon after, it was applied in corporate settings as a managerial selection tool made popular by the results of the Management Progress Study at AT&T (Bray & Grant, 1966). Assessment centers use situational exercises (e.g., leaderless group discussion, in-basket, oral presentation) to simulate managerial jobs. While participating in the exercises, student performance is evaluated by trained observers based on demonstrations of managerial skills (e.g., oral communication, teamwork, initiative).

Thus, the assessment center provides a realistic method of evaluating assessee performance in situations similar to those encountered by managers on the job—i.e., "authentic assessment." There is evidence suggesting that ACs are broadly used by organizations employing business school graduates (Spychalski, Quinones, Gaugler, & Pohley, 1997). In addition, research on ACs has established the methodology as a strong predictor of managerial success (Howard, 1997; Gaugler, Rosenthal, Thornton & Bentson, 1987).

Using ACs in an academic setting is not an altogether novel idea. In 1985, the AACSB in partnership with Developmental Dimensions International (DDI) employed AC technology for evaluating student managerial skills. This program, however, was discontinued after a few years due to financial burdens. However, a number of universities have successfully implemented AC technology (e.g., Alverno College, Arizona State University, California State-Fullerton, Case Western Reserve University, Central Missouri State University, Claremont-McKenna College, Indiana University, Texas Tech University, Southern Illinois University-Edwardsville, Valparaiso[32]) for purposes of skill development and/or outcomes assessment. Many of these assessment

centers have been discussed in the literature and during recent conference presentations (e.g., Riggio, *et al.*, 2003).

Research on academic ACs has shown that the skills assessed and developed via AC methods are truly important for future student success. For example, Waldman and Cullen (2001) described an academic AC in which student performance was significantly related to post-college salary increases and promotions. Schleicher, Riggio and Mayes (2001) found that job performance could be predicted by assessment center performance five years post graduation. Riggio *et al.* (2003) found that 42% of the variance in rated "employability" was due to interpersonal skills, such as oral communication and leadership, measured in their assessment center. These research efforts demonstrate that academic ACs are hitting the mark with respect to the types of skills being assessed. That is, academic ACs measure important, job-related behaviors that are highly related to graduates' immediate and long-term job performance.

360 Degree Feedback or Multi-Source Feedback Assessment

Another innovative method used to assess interpersonal skills is 360-degree feedback. The term "360 degree feedback" comes from the idea that raters "in a full circle" around the target are involved with providing their input. The most common example of this model can be seen when an employee is rated by his or her boss, a number of peers, all of the people who are direct reports, and possibly even customers, clients, or other people with whom the target has frequent interaction. All of these sources of input are in addition to the self-evaluation that is provided by the person who is undergoing the 360-degree assessment.

The rise of 360-degree feedback lies in the roots of the human relations movement of the 1950s and 1960s and is an offshoot of the survey/feedback approach used at the organizational level. Thus, while traditional survey/feedback has been used to provide feedback regarding organizational processes and employee responses across large groups, 360-degree feedback programs are specifically targeted toward "feeding back" information to specific individuals (usually supervisory or managerial level employees) about their work behaviors.

While we are not aware of any specific research that has focused upon 360-degree feedback in purely academic settings, the practice is certainly not alien to the classroom. Studies have been conducted using working MBA students (Brett & Atwater, 2001), but the general idea of most of the research has been that the MBA students actually represent a varied population of organizational members. Thus, working MBA students are generally of more interest when it comes to research, because they are not "typical students." The use of multi-rater feedback, however, is common in everything from evaluating team presentations to class participation.

Assessment Models

In our experience, ACs and 360s can be used at different points in the curriculum to serve different objectives, making them flexible techniques that address a series of assessment and curricular issues. To provide a better sense of how these assessments can be used, we will present two common scenarios with which we have been associated.

Program- Focused Assessment Model

The most common approval is what we will call a *program-assessment* model.[33] Using this program-assessment approach, students are put through an assessment center or a 360-degree assessment designed to measure a targeted set of behavioral skills (e.g., leadership, decision-making, communication, teamwork, etc.) somewhere near the beginning of their business school experience. Due to practical issues (e.g., scheduling, credit hour assignment, attendance, etc.), students normally complete the assessment activity as a part of a required course (e.g., Principles of Management) that comes near the beginning of the business school curriculum. In MBA programs, this assessment tends to take place as part of the orientation or during the program's first semester, allowing early feedback to the student regarding his/her skill level. In order to serve as a value-added program assessment, this initial pre-test is followed up with a post-test that occurs near the end of the curriculum.[34] In undergraduate and MBA programs alike, capstone strategy courses tend to be taken by students very near the end of the program. Thus, we have found capstone courses to be practical and relevant places to capture this post-test assessment.

A common alternative in conducting program assessment is through the use of a post-test measurement alone. In this approach, there is no pre-test to assess where the students "start" in their skill-building, only a measure of where they "finish." So, much like the post-test portion of the technique described above, students near the end of their program (often in a capstone strategy course) go through an assessment to provide data used to measure their skill proficiency. This approach is useful to the students as well as the program. For the student, it provides an assessment of specific skills that they may need to develop further once they graduated.

Student-Focused Assessment Model

Another popular approach takes place within the context of a specific course. For example, a number of schools have required courses that are focused specifically on "skill development." In contrast to the program assessment model described above, these schools often utilize a *student-focused assessment* approach. In this approach, students complete an assessment activity (normally an assessment center, but occasionally a 360-degree assessment) at the beginning of a term (obviously, the dimensions used in the assessment activity are specific to the content of the skill

111

development course) and then receive feedback within the first few weeks of the course. This way, students are informed of their individual performance and thus have an individual benchmark against which to judge their future performance. Since the skill development course is designed to provide specific developmental opportunities, a post-test needs to occur near the end of the term so that students can be re-evaluated.

Curriculum Example: Assessment Centers

As discussed above, we have employed our decision rules regarding what skills to assess/include in an assessment center effort for Indiana University. As a result of this process, we have identified seven specific interpersonal skills dimensions we believe meet our criteria. These are broad categories of skills, each containing very specific, smaller skills. The seven skill dimensions assessed in the assessment center are: decision-making, initiative, leadership, planning, organizing, teamwork and oral communication. Certainly other important skill dimensions exist, but our experience has shown us that these seven are quite generalizable and provide an excellent starting point.

When designing an assessment center, resources and numbers of students to be assessed need to be seriously considered. Relevant factors to consider are the skills to be assessed and the significant logistics involved with some approaches. When we designed the Indiana assessment center, we needed to be able to simultaneously assess at least 40 students, and under special situations we knew that we needed to be able to assess 80 students simultaneously (in some cases, we have assessed 1,000 students in 3 days). Because of the large numbers of students, we designed the assessment center with the following guidelines:

1. All assessment activities would need to be self-contained. In other words, students would need to be able to work on activities without a facilitator being part of the activity.
2. The assessment could not last longer than 3 hours due to scheduling, space, and student fatigue considerations.
3. Costs needed to be contained. The cost structure needed to be affordable for students and/or the college without sacrificing quality exercises and assessment.
4. To reduce costs and increase accuracy, all rating would be done via videotape (versus live rating as is often done in ACs) so that rating could occur more efficiently (using fewer total raters) and more reliably (a permanent record would exist to be reviewed for accuracy).

Once the skills to be assessed have been identified, exercises need to be developed that provide students the opportunity to demonstrate the skills being assessed. In most ACs, participants engage in a number of activities

112

designed to elicit their skills, including leaderless group discussions, presentations, in-baskets, and role-playing. We have designed four specific exercises: two leaderless group discussions, an oral presentation and an in-basket exercise. One leaderless group discussion (LGD) requires students to decide on their top three candidates (out of a pool of seven) for a new CEO position by evaluating candidate resumes. The other LGD instructs participants to solve a customer service issue impacting the organization. The leaderless group discussions typically have five or six members. The oral presentation requires students to deliver a three-minute speech in front of a small audience of their colleagues (three or four other students) on an organizationally relevant topic. Finally, the in-basket exercise entails multiple memoranda and organizational information that require written responses. Although many ACs incorporate one-on-one activities that involve the assessee interacting with a role-player,[35] this type of activity has been deemed too labor intensive for the large-scale on which we were operate. More specifically, because of the number of students in Indiana University's Kelley School of Business, we calculate that we would need 26 role-players working for the better part of three straight days to have a well-run, one-on-one exercise. Faced with the problems of trying to keep consistency across this number of role-players (both between the role players and across individual role players as they fatigued), we have decided that a one-on-one activity is not viable.

In addition to the exercises described above, our exercises also include self-assessments. Students complete motivational (e.g., need for achievement, affiliation and power) and personality assessments (e.g., The Big Five), as well as self-ratings on the skill dimensions. These additional measures are distinct from the skills feedback, and they help students to put their skills feedback into a larger framework and learn about how their preferences impact their own and others' behavior.

To provide a richer assessment experience, we have chosen to integrate the four exercises into a compressed workday (the actual assessment takes 2 1/2 hours) for a simulated company. At least one week before the assessment, students receive the company's annual report (the company is referred to as Iliad, Inc. and the assessment center is referred to as "Iliad"). This annual report provides important company background (basic financials, letter from the CEO, etc.) and provides context for the students. Building on this initial context, when the students arrive at Iliad they are shown a direction video and then provided with their personalized in-basket. The in-basket is the exercise that integrates the other activities into the assessment experience. In others words, the tasks are integrated into students' schedules so that memos from the in-basket direct the students to their meetings and to their presentation. In-basket memos are to be answered when the students are not involved in the other activities. By doing this, the assessment center maintains high face validity and fidelity, because it has the feel of a "day at work."

Sample Assessment Center Rating Tools

Because the four exercises (2 leaderless group discussions, presentation, and in-basket) associated with Iliad are designed around a series of seven skill dimensions, it is important for validity purposes to have multiple opportunities during the assessment for the student to exhibit each skill area. For instance, providing useful feedback on a behavior (e.g., oral communication) that was only observed for a few minutes could provide very unreliable or unrepresentative results and provide subsequently wrong or biased feedback. As a result, all skills are assessed in multiple activities so that students have multiple opportunities to demonstrate the skills. More rating opportunities tend to provide more stable and accurate ratings of the skill. The table below provides the skill/exercise matrix on which Iliad operates.

Table 2
Iliad Skill/Exercise Matrix

Skill/Exercise	New CEO Meeting	Customer Service Meeting	3-Minute Presentation	In-Basket Memos
Decision-Making	X	X	X	X
Initiative	X	X	X	X
Leadership	X	X		
Planning	X	X	X	X
Organizing	X	X	X	X
Teamwork	X	X		
Oral Communication	X	X	X	
Note: X denotes that the skill is assessed in the respective exercise				

In order to facilitate objectivity in the rating process, we give great deference to the rich empirical literature on designing ACs (e.g., Lievens, 1998). A number of suggestions follow from that literature. First, we have created a very precise behavioral explanations or "rating dictionary" where each behavior is defined for each meeting. Further, we choose to use a behavioral checklist rather than a Likert-type rating scale (Reilly, Henry & Smither, 1990). That is, raters assess skills using a binary system noting whether the student exhibited the behavior or not. For instance, in the two leaderless group discussions (i.e., the new CEO meeting and the customer service meeting), teamwork behaviors are rated. While they are assessing the meetings, raters are looking for a series of specific behaviors that represent the dimension of teamwork. Some examples of these specific behaviors include:

• Seeks input from other group members

- Validates other group members
- Does not interrupt other group members
- Checks for common understanding among group members

To further elaborate on the teamwork behaviors listed above, the "rating dictionary" provides a further description of the behavior and provides prototypical examples.

Example-Iliad Rating Dictionary Excerpt:

Selected Teamwork Behaviors Rated in New CEO Meeting
Behavior: Seeks input from other group members.
Definition: Subject asks questions for clarification (e.g., "Can you provide a further example of that?") or solicits ideas from others (e.g., "We haven't heard your thoughts yet, what do you think about this candidate?").

Behavior: Validates other group members.
Definition: Affirms others and/or their contribution ("I thought her idea was really good." This is NOT simply nodding in agreement or saying "right" or "I agree"). Exemplary demonstration of this behavior is shown when the subject both affirms the contribution of another person and then "piggybacks" on their idea (e.g.," I'd like to take that one step further…").

Behavior: Does not interrupt other group members.
Definition: Does not interrupt other group member inappropriately (e.g., talking over or "cutting off" another person who is making a concise point.) This behavior should still be checked if the ratee stops a "rambler" or apologizes for the interruption and provides a reason (e.g.," I am sorry to interrupt, but we need to make a decision in the next two minutes").

Behavior: Checks for common understanding among group members.
Definition: Verbally confirms the positions and/or ideas expressed by group members. Common examples include: "So, it seems that we have decided upon…, does everyone agree?" or "In summary, I hear…, is this correct?" or "It sounds like everyone is on the same page, does everyone support this answer?"

These rating dictionaries, however, have to include all of the individual behaviors representing the overall behavioral dimension (in the example it was teamwork) for each exercise (in the example it was the new CEO meeting). The use of a detailed rating dictionary will increase the reliability of the ratings by substantially reducing the amount of judgment that is left to

115

the individual raters. This increased reliability is also important, because it makes comparing the ratings of two different raters relatively trouble free as well as making comparisons between schools possible.

Once all of the ratings are compiled, they are entered into a scoring program, where each observed behavior (e.g., seeks input from other group members) is provided a weight. Subject matter experts (50 practicing managers) have rated how important the behaviors are for successful managers. These ratings are used to determine the relative weight of the observed behaviors. The weighted scores of the observed behavior are then summed to provide a score by behavioral dimension (e.g., teamwork) and exercise (e.g., CEO meeting). These data are then aggregated across activities (e.g., since teamwork is measured in the two group meetings only, the teamwork score is the total points earned in these two meetings) and across behaviors within an exercise (e.g., a new CEO meeting score is derived by summing the points earned in each of the seven dimensions, since all seven dimensions are assessed in the new CEO meeting). This way, feedback is based upon both the student's skills across activities and their performance within a specific exercise.

Sample Student Feedback

Through considerable experience, we have found that students process their feedback best when it is presented by both skill (e.g., teamwork score, planning score, etc.) and exercise (score in new CEO meeting, presentation, etc.). We have also found that providing this information in a benchmarking format (i.e., by percentile) provides the student excellent information on how they performed relative to thousands of other people who completed the same assessment under the same conditions. Because the scores are being compared to a large database of other scores, we are able to present scores as percentiles. Some institutions like to provide their students the actual percentile scores, whereas other schools prefer to group bands of scores under descriptive labels (e.g., the bottom 25% = "needs improvement," middle 50% = "average," top 25% = "outstanding"). The framing of the feedback is largely based on the institution's assessment and development goals and varies by institution and professor. For instance, some schools like to "light a fire" under their students, providing motivation for performance improvement. These schools tend to prefer a raw percentile score shown relative to others nationally. Other schools tend to use more descriptive terms and give more general feedback to students. To provide a flavor of the numerical feedback provided to students, we have included an excerpt from the report each student receives. The entire report can run about 15 pages long and includes a developmental action planning process. In Appendix A, we have provided an example where both percentile and descriptive labels are used.

Some AC Challenges

In our experience with ACs, we have experienced a number of challenges surrounding issues of design and administration. First, it is difficult to fully express the enormous time commitment associated with designing an effective and sustainable AC. This is not a project where a faculty member can be given a course release and be expected to have a "well-oiled" AC at the end of a term. For the first three years, we made significant changes after almost every major group assessed, and WE had the benefit of running the AC at multiple institutions. This allowed for significant improvements to occur much more quickly. Second, the design of good exercises that truly elicit the desired assessed behaviors is difficult; expect many iterations and "versions." Third, developing checklists of truly observable behaviors is much more difficult than it may appear at first. For instance, in the early development stages of Iliad, we were convinced that active listening was an important skill that needed to be included in the rating. However, after having a large number of people watch hours of videotape, we came to the conclusion that active listening could not be rated to our satisfaction. In all, the design aspect of a good assessment center needs to be considered in months and years, not days and weeks.

Beyond the design challenges, the administrative challenges can be surprisingly complicated even with a supportive administration and faculty. The general challenges from an administrative end relate to rating, physical space and time, funding, and curriculum integration. Regardless of whether the assessment center is rated live or via videotape, trained raters are needed. The selection, training, rating system design, rater compensation, and data management issues all have to be well thought out in advance. In the end, the information generated from any assessment center is only as good as the raters and the rating system they employ.

The requirements of both physical space and student time are interrelated issues that often catch programs off guard. Due to the use of multiple exercises, the number of separate rooms required can grow quite large. For instance, we generally utilize about seven rooms when forty people are being evaluated. In most institutions, this amount of space is a scarce resource during the regular term schedule. In addition, most ACs require more student time than a regular class period allows. As a result of this problem, we kept Iliad under three hours so that it could take the place of one week of classes. Another solution to space and time concerns is to assess students on weekends, so that space constraints are relieved and student schedules are not as restricted. Obviously, however, coming in for three hours on a Saturday or Sunday morning will not be universally well received by students and even participating faculty.

The third administrative issue is how to fund an ongoing AC. Our experience with different institutions shows there are multiple ways to find fund the AC effort; however, the funding should be commensurate with the

goals. That is, if the purpose of the AC follows the program-focused assessment model, it makes most sense for administration to provide funding. In a course-focused assessment model, institutions usually choose to have the students cover the cost of the AC through course fees or course materials such as a textbook ancillary. Each of these is a viable option, but a system does need to be put in place that can handle a recurring expense stream. Because of the recurring nature of the expense, it is our experience that funding through one-time grants, etc. is not a viable long-term approach.

The final administrative challenge that we will mention here is that follow-up to the AC – what assessment experts refer to as "closing the loop"—is critical. At a program level, the information from the AC should flow back into the decision-making process and be used to modify the curriculum (e.g., students are not demonstrating teamwork skills, so we need to focus more on these skills). At the course level, the identification of weaknesses in students raises the expectation that the institution will be doing something (i.e., providing a class or some other developmental activity) to help students develop their skills. Many business schools do not have such developmental experiences in place at present.

360-Degree Assessment Example

In contrast to the AC example provided above, it is much easier to modify 360-degree assessments for specific instances. The "start-up" costs of modifying a survey (associated with the 360 assessment) are substantially less than those associated with modifying assessment center activities. On the other hand, 360s can be difficult to achieve due the need for a large (e.g., 8-20 are desirable) group of raters. Thus, 360s are most viable when a large portion of the assessees are currently employed (e.g., part-time MBAs or non-traditional undergraduates), recently employed (e.g., full-time MBAs), or are engaged in other settings where their skills are displayed in the presence of other people (e.g., substantial group interaction, sports teams, extracurricular activities, etc.). Moreover, we have generally used 360s to assess the types of skills that may be difficult to evaluate in an AC, such as those that are more complex and require longer observation periods. As a result, we have used 360-degree assessments to examine skill dimensions such as: develops others, motivates others, networking behavior, administrative skills and persuasiveness.

To provide a detailed example of a 360-degree assessment survey, we include a sample instrument and explain the process. To facilitate a cost effective technique for gathering and processing 360-degree assessments for a large number of students, we have developed a Web-based system through which the data can be collected and feedback provided. The first step in conducting any assessment (and the 360-degree assessment is no different) is to determine what skills are going to be assessed. Based on the goals of the program with which we have been working, they have identified

118

the following seven areas. As a result, the skills to be assessed are: Administrative Skills, Communication Skills, Interpersonal Skills, Leadership & Coaching Skills, Political Skills, Motivational Skills, and Service Skills.

Once the skill dimensions were identified, specific behavioral survey items needed to be developed which tapped these underlying skill dimensions. After working with the instructors to better understand their specific operationalizations of the dimensions, we constructed a 42-item survey designed to measure the seven skill areas. This instrument is rated on a 5-point Likert scale asking the rater to assess level of agreement with the statement. In addition, raters provide written comments about the student's strengths and weaknesses. The survey requires about 20-30 minutes to complete. Sample items from this survey are included in the table below:

Because we use an online system, the first step in the process is getting the e-mail addresses of all of the students to be assessed. The

Table 3
Sample 360 Items

Administrative Skills
• Delegates responsibilities appropriately • Manages meetings effectively
Communication Skills
• Speaks clearly in front of groups • Conveys information clearly in written documents
Interpersonal Skills
• Works with others to effectively resolve conflicts • Develops cooperative working relationships
Leadership & Coaching Skills
• Leads by example • Provides specific constructive feedback in a timely manner
Political Skills
• Understands the agendas and perspectives of others • Recognizes key stakeholders related to important decisions
Motivational Skills
• Persists in the face of obstacles • Establishes challenging goals
Service Skills
• Anticipates customers' needs • Shows a concern for customer satisfaction

students being assessed then receive an e-mail asking them to provide the e-mail addresses of people who should be contacted as raters (e.g., superiors, co-workers, subordinates, customers, etc.). Once the rater list is entered, each rater receives an e-mail including a link that explains the purpose of the assessment and also gives the rater an opportunity to provide their assessment of the student being assessed. In addition, the student being assessed receives an e-mail with a link allowing him/her to complete their self-assessment to "round out" the 360 process.

A critical component of the process involves explaining to raters the purpose of the 360. Most important is to ensure the confidentiality and anonymity of the raters. Although it is beyond the scope of this chapter, it should be noted that 360s can be quite harmful if not handled ethically and professionally (see Waldman, Atwater, & Antonioni, 1998; DeNisi & Kluger, 2000). Since the number of raters responding impacts the quality of the feedback, it is imperative that the raters understand the voluntary nature of the process, why their voice is critical towards student development, and how the data are used. The process is highly time-sensitive, so raters often need a few reminders to ensure a turn-around time.

Once the surveying portion of the 360-process has been closed, students are able to access their online feedback reports. We have customized these feedback reports based on a number of different possibilities, but generally follow a format that moves from broad to specific feedback. As a result, we start with the feedback provided at the level of the skills dimensions and compare the self-rated scores to the scores provided by all of the other raters. Then, we "break out" the feedback by rater category (e.g., superiors, peers, subordinates, etc.). In subsequent report sections, we provide the feedback by item (in the example, all 42 behaviors) in the same manner (i.e., comparison of self vs. others and then by rater category). Lastly, verbatim comments are provided from all of the raters. After all of the feedback is presented, a developmental action planning guide is also included to help the student interpret the data and to provide a structure for making future improvements.

Although the entire report is relatively long (about 12 pages), we have excerpted a feedback report and have provided an example in Appendix B. In addition to the information provided in Appendix B, a full feedback report also includes an analysis by item and a significant amount of information to guide the assessee through the report.

Some 360 Challenges

Much like ACs, 360-degree assessments can be considered to have both design and administrative challenges, although the relative difficulty of the challenges faced differ significantly. From a design perspective, the actual 360 instrument is not particularly difficult to design once the skills to be assessed have been identified. A few faculty members with particular expertise in survey design and knowledge of management skills can typically develop

a competent 360. In this way, 360s offer much less of a design burden than do ACs. On the other hand, the design of the mechanism that allows for the surveys to be distributed, collected, and tabulated can be extremely labor intensive. We generally perform 360s online for the obvious benefits seen in reduced administrative burdens. This process can be managed on paper, but the data entry, computations, and summarization can be very labor intensive (especially if the group of students is relatively large). Next is the availability of useful 360 raters. Normally, we conduct 360-degree assessments with nontraditional students who have full-time employment. Conducting a 360-degree assessment for an individual who does not currently work, or who has recently held a job, can be a significant problem because of the lack of appropriate raters (i.e., people who can provide informed judgments regarding the individual's behavior). In some cases this problem can be alleviated if the curriculum includes significant, intensive, teamwork where students have a great deal of interaction with others who can provide ratings. If, however, the student does not have current or recent work experience and does not interact with other students on a regular basis, 360-degree assessments are not appropriate.

From an administrative viewpoint, 360-degree assessments offer some of the same challenges as do ACs. More specifically, both funding and curriculum integration are issues that must be addressed. Most universities choose 360s for use in a course or student-focused model. Thus, funding is primarily the responsibility of the student. Like the AC, follow-up from 360s is critical if students are to truly develop the skills assessed. Ideally, the feedback will be provided in an individual feedback session by a trained development coach. The coach can assist the student in processing the feedback and creating an action plan. The likelihood of this occurring is negatively correlated with class size. That is, it may be impossible to meet individually with students from a large class. There are several potential places to look for coaches— professors, advisors, career center staff, executives in residence, etc.

Concluding Thoughts
We do not intend to overstate the value of assessment centers and 360-degree feedback in their ability to assess and develop interpersonal skills. In fact, some scholars are highly critical of both ACs (e.g., Sackett & Dreher, 1982) and 360s (e.g., Waldman, et al., 1998). However, in examining alternatives, it is difficult to find other methods that are able to more reliably and systematically measure behavioral skills and have higher face validity. ACs and 360s offer both an effective and efficient means of systematically assessing interpersonal skills on the way towards student skill development. Although the administrative obstacles associated with these assessment procedures may seem insurmountable, with a long-term commitment, they can be overcome. The result of such commitment is an effective and sustainable approach to assessing and developing managerial skills.

Appendix A
Iliad Assessment Center Results for Joe Sample

Student ID # 06122
Iliad ID# S35

These results are being provided to you for your development. These results should always be considered in a larger context. In this case, look at the results of this assessment, and consider other feedback you have received in the past. Look for trends in the information, as these will provide you with the most useful and true information.

Detailed Explanation of Your Assessment Center Performance by Skill Area

Below are details about your performance for each skill area. The table summarizes your scores for each skill area. Under the table are examples of the types of behaviors that were assessed to arrive at your score. The percentile ranking is the percentage of others that you outperformed on this skill (i.e., higher numbers are better).

Assessed Skill	Percentile	Description
Decision-Making	87	Excellent
Initiative	71	Very Good
Leadership	55	Good
Planning	2	Needs Improvement
Organizing	25	Needs Improvement
Teamwork	14	Needs Improvement
Oral Communication	32	Good

In the next portion of the feedback, their actual performance in the assessment center is compared to their skill self-assessments that were collected immediately before the assessment began. The feedback from the combination of their skill self-assessment and their assessment center performance is provided in the following manner:

122

Iliad Assessment Center Results for Joe Sample (Part 2)

Student ID # 06122
Iliad ID# S35

Summary of Your Assessment Center Skill Areas
Summary Explanation

At your assessment, you provided a self -assessment of your skills in the following areas. Those scores were compiled and any score of over 50 was considered a self-rated strength and any score 50 or below was considered a self-rated weakness. Then, your actual performance was compared to the self-ratings. The table below includes the information that is a summary of your performance versus the self-ratings on the following skills.

- Decision-Making, Initiative, Leadership, Planning, Organizing, Teamwork, Oral Communication

Unacknowledged Strengths are skill areas in which the assessors rated you above average (compared to other participants) in this area, but you rated yourself lower than other participants. These are skill areas you may not be aware you possess.

Confirmed Strengths are skill areas in which the assessors rated you above average and you also rated yourself above average. These represent skill areas on which you can capitalize.

Confirmed Weaknesses are skill areas in which the assessors rated you below average and you also rated yourself below average. These are skills you already recognize have room for improvement.

Unacknowledged Weaknesses are skill areas in which the assessors rated you below average and you rated yourself above average. These represent skill areas that you may not recognize need improvement and therefore they are a good place to begin individual development.

Skill Assessment Grid

Skill Areas	Confirmed	Unacknowledged
Strength	Decision-Making, Initiative	Leadership
Weakness	Planning, Organizing	Teamwork, Oral communication

The third portion of the feedback provides some information relating to the student's performance in each of the activities. While most assessment centers focus solely upon the skill-related feedback, we found that many students wanted to know, "How did I do on the speech?" or "How was my

CEO meeting?' To accommodate these requests, we decided that adding performance feedback based upon the activities was consistent with the spirit of developmental feedback and may be more meaningful to certain students. An example of this feedback by exercise is presented below:

Iliad Assessment Center Results for Joe Sample (Part 3)

Student ID # 06141
Iliad ID# S35

Your Assessment Center Performance Score by Exercise
Recall that you completed four exercises during the assessment center:

- CEO Selection Meeting
- Customer Service Meeting
- Oral Presentation
- In-basket Exercise

Below are your scores for each assessment center exercise. When reading the table below, please remember that the scores from one exercise (e.g., the CEO selection meeting) are independent of the other activities (e.g., the in-basket). Because of this, a "higher" score on your speech than in your Customer Service meeting does not mean a high performance level. The performance level is based upon the percent of your classmates that you did better than. The same is true of your overall score - this is ranked against all of the other total scores to come up with your overall performance level.

Exercise Name	Percentile	Description
CEO Selection Meeting	42	Average
Customer Service Meeting	38	Good
Oral Presentation	70	Very Good
In-Basket	58	Good
Total	56	Good

Appendix B

Excerpt to Illustrate a 360 Feedback Report for Jen Sample

Self and All Observers Specific Behaviors Table

Specific Behaviors	Self-Rated Score	All Others-Rated Score
Administrative Skills	4.86	3.99
Communication Skills	4.40	3.88
Interpersonal Skills	3.57	3.71
Leadership & Coaching Skills	3.71	3.92
Political Skills	2.50	3.81
Motivational Skills	4.43	4.23
Service Skills	3.80	3.94

Self and All Observers Specific Behaviors Table

Specific Behaviors	Self	Boss	Reports	Peers	Others
Administrative Skills	4.86	4.33	3.70	4.11	3.79
Communication Skills	4.40	4.10	3.80	3.92	3.70
Interpersonal Skills	3.57	3.92	3.56	3.69	3.79
Leadership & Coaching Skills	3.71	4.12	4.00	3.80	3.93
Political Skills	2.50	4.00	3.50	3.88	3.88
Motivational Skills	4.43	4.64	3.90	4.26	4.21
Service Skills	3.80	4.20	3.93	3.81	4.00
Note: A full report includes detailed analyses by item and substantial information designed to help the assessee interpret the results.					

125

Endnotes

32 The Assessment Center at Valparaiso in described in detail in "Fostering the Professional Development of Every Business Student: The Valparaiso University College of Business Administration Assessment Center," in Vol. 1, No. 2.

33 Program assessment models can be either internally developed, or contracted out. The Valparaiso example in volume 2 of this series is an example of an internally developed center. Quite often, universities need external help in this process as the demands are often quite technical and the learning curve to constructing these processes is quite steep. For information on the assessment center that we have helped universities institute over the past ten years, please contact the authors.

34 Editor's note: The AACSB does not require value-added assessment, and does not specify when assessment must occur. Thus, the AC would not have to be administered as a pre-posttest, or at the end of the curriculum, in order for it to be deemed an acceptable assessment.

35 Editor's note: Valparaiso's assessment center incorporates role play—and, yes, it is a labor intensive process.

References

Anderson, J. R. (1985). *Cognitive psychology and its implications*. (2nd ed). New York: Freeman.

Bartels, L.K., Bommer, W.H. & Rubin, R.S., (2000). Student performance: Assessment centers versus traditional classroom evaluation techniques. *Journal of Education for Business, 75*(4), 198-201.

Bigelow, J.D. (1995). Teaching managerial skills: A critique and future directions. *Journal of Management Education, 19,* 305-325.

Bigelow, J., Seltzer, J., van Buskirk, W., Hall, J., Schor, S. Garcia, J. & Keleman, K. (1999). Management Skills in Action: Four Teaching Models. *Journal of Management Education. 23* (4), 355-376 Boyatzis, R.E., Stubbs, E.C. and Taylor,

Boyatzis, R.E., Stubbs, E.C. and Taylor, S.N. (2002). Learning cognitive and emotional intelligence competencies through graduate management education. *Academy of Management Journal on Learning and Education. 1*(2), 150-162.

Bray, D. W., & Grant, D. L. (1966). The assessment center in the measurement of potential for business management. *Psychological Monographs, 80*(17), 1-27.

Brett, J. F. & Atwater, L. E. (2001). 360° feedback: Accuracy, reactions and perceptions of usefulness. *Journal of Applied Psychology, 86,* 930-943.

DeNisi, A. S. & Kluger, A. N. (2000). Feedback effectiveness: Can 360-degree appraisals be improved? *Academy of Management Executive, 14*(1), 129-139.

Eberhardt, B.J., McGee, P., & Moser, S. (1997). Business concerns regarding MBA education: Effects on recruiting. *Journal of Education for Business, 72*(5), 293-296.

Gaugler, B.B., Rosenthal, D. B., Thornton, G.C., & Bentson, C. (1987). Meta-analysis of assessment center validity. *Journal of Applied Psychology, 72*(3), 493-511.

Harris, M. M. & Schaubroeck, J. (1988). A meta-analysis of self-supervisor, self-peer, and peer-supervisor ratings. *Journal of Applied Psychology, 41*, 43-62.

Howard, A. (1997). A reassessment of assessment centers: Challenges for the 21st Century. *Journal of Social Behavior and Personality, 12,* 13-52.

Huselid, M.A. (1995). The impact of human resource management practices on turnover, productivity, and corporate financial performance. *Academy of Management Journal, 38,* 635-672.

Judge, T. A. & Piccolo, R. F. (2004). Transformational and transactional leadership: A meta-analytic test of their relative validity. *Journal of Applied Psychology, 89*(5), 755-768.

Kraiger, K., Ford, J.K. & Salas, E. (1993). Application of cognitive, skill-based and affective theories of learning outcomes to new methods of training evaluation. *Journal of Applied Psychology, 78,* 311-328.

Lievens, F. (1998). Factors which improve the construct validity of assessment centers: A review. *International Journal of Selection and Assessment, 6,* 141-152.

McConnell, R.V. & Seybolt (1991). Assessment center technology: One approach for integrating and assessing management skills in the business school curriculum. In J.D. Bigelow (Ed.), *Managerial skills: Explorations in practical knowledge.* Newbury Park, CA: Sage.

Mintzberg, H. (1975). The manager's job: Folklore and fact. *Harvard Business Review, 53,* 49-71.

Pfeffer, J. & Fong, C. T. (2002). The end of business schools? Less success than meets the eye. *Academy of Management Learning and Education, 1,* 78-96.

Porter, L. & McKibbon, L. (1988). *Management education and development: Drift or thrust into the 21st century.* New York: McGraw-Hill.

Reilly, R.R., Henry, S. & Smither, J.W. (1990). An examination of the effects of using behavior checklists on the construct validity of assessment center dimensions. *Personnel Psychology, 43,* 71-84.

Riggio, R.E., & Riggio, H.R. (2001). Self-report measurement of interpersonal sensitivity. In J.A. Hall & F. Bernieri (Eds.), *Interpersonal sensitivity: Theory and measurement.* (pp. 127-142). Mahwah, NJ: Lawrence Erlbaum Associates.

Riggio, R. E., Mayes, B. T., & Schleicher, D. J. (2003). Using assessment center methods for measuring undergraduate business student outcomes. *Journal of Management Inquiry, 12,* 68-79.

Rynes, S. L., Trank, C. Q., Lawson, A. M., Ilies, R. (2003). Behavioral coursework in business education: Growing evidence of a legitimacy Crisis. *Academy of Management Learning & Education, 2,* 269-274.

Rynes, S. L. & Trank, C. Q. (1999). Who moved our cheese? Reclaiming professionalism in business education. *Academy of Management Learning & Education, 2,* 189-206.

Sackett, P.R., & Dreher, G. F. (1982). Constructs and assessment center dimensions: Some troubling empirical findings. *Journal of Applied Psychology, 67,* 401-410.

Schleicher, D. J., Riggio, R. E., and Mayes, B. T. (2001). New frame for frame-of-reference training: Enhancing the construct validity of assessment centers. *Journal of Applied Psychology, 87,* 735- 747.

Spychalski, A. C., Quiñones, M. A., Gaugler, B. B., & Pohley, K. (1997). A survey of assessment center practices in organizations in the United States. *Personnel Psychology, Vol. 50,* pp. 71-90.

Thompson, K. R. (2004). A conversation with Milton Blood: The new AACSB Standards. *Academy of Management Learning & Education, 3,* 429-440.

Waldman, D. A., Atwater, L. E., & Antonioni, D. (1998). Has 360 feedback gone amok? *Academy of Management Executive, 12,* 86-94.

Waldman, D.A. & Cullen, T. (2001). *Using academic assessment center data in the prediction of early career success.* Paper presented at the Annual meeting of the Society for Industrial and Organizational Psychology, San Diego, CA.

Waldman, D. A., & Korbar, T. (2003). Student assessment center performance in the prediction of early career success. *Academy of Management Learning & Education, 3,* 151-168.

Waters, J. A., Adler, N.J., Hartwick, J. & Poupart, R. (1983). Assessing managerial skills through a behavioral exam. *EXCHANGE: The Organizational Behavior Teaching Journal, 8,* (2), 37-44.

Whetten, D. A. & Cameron, K. S. (1995). *Developing management skills* (3rd ed.). New York: HarperCollins.

Authors Bios

William H. Bommer (*w.bommer@csuohio.edu*) is an assistant professor in the management and labor relations department at Cleveland State University's Nance College of Business. He received his Ph.D. in Organizational Behavior from Indiana University. His research interests include transformational/transactional leadership, organizational citizenship behavior, leadership development, and research methods.

Robert S. Rubin (*rrubin@depaul.edu*) is an assistant professor in the management department at DePaul University's Kellstadt Graduate School of Business. He received his Ph.D. in Organizational Psychology from Saint Louis University. His current research interests include transformational leadership, assessment centers, social and emotional individual differences, and management development.

Lynn K. Bartels (lbartel@siue.edu) is an associate professor of Psychology at Southern Illinois University-Edwardsville. She received her doctorate in Industrial/Organizational Psychology from the University of Akron. Her research interests include assessment centers, interviews and employee selection, and development.

One of the important questions faculty must ponder when choosing learning goals for their curriculum is "how do we expect our students, as a result of our program, to think?" Often the answer to this question is "critically." There is, however, a lack of understanding of what that term really means. In this chapter, a respected scholar on critical thinking and its role in college curriculum provides valuable insight on how to define, assess, and cultivate this important (and popular) learning goal.

CHAPTER 8
ASSESSMENT OF CRITICAL THINKING
Susan K. Walcott
WolcottLynch Associates

An Introduction to Critical Thinking

Most colleges and degree programs include critical thinking as a learning goal. Educators view critical thinking as essential for their students' professional, personal, and civic success. Although there is general agreement that critical thinking is important, disagreement often exists about how it is defined. Furthermore, no widely accepted instrument exists for assessing student critical thinking skills. Given these difficulties, many educators are uncertain about how they should assess critical thinking in their courses. Departments, programs, and colleges face similar problems when asked to address critical thinking in their formal assessment plans.

However, there are several reasons why faculty and programs should assess critical thinking. First, programs are often under a mandate from accrediting organizations, legislatures, or others to assess stated learning outcomes, which include critical thinking. Second, programs and courses that cite critical thinking as a learning goal are accountable to their students and other stakeholders for assessing that goal (e.g., Black & Duhon, 2003). Third, critical thinking affects student performance on many existing course assignments, and faculty members can do a better job of grading and providing students with appropriate feedback if they more explicitly assess critical thinking skills. Fourth, critical thinking assessment can lead to new understandings of student skills and the effects of educational activities, which in turn can lead to improved student learning. Finally, although no absolute answers exist in the higher education literature about how best to assess critical thinking, considerable guidance does exist. It is possible to implement reasonable methods for assessing critical thinking in programs and individual courses.

Defining Critical Thinking Skills

Before any type of learning outcome can be assessed, it is first necessary to clearly define the desired learning outcome. Some programs and professors have already defined their critical thinking skills. However, others have not adequately articulated what they mean by critical thinking, nor have they described the critical thinking skills that students must exhibit. This section explores a range of resources that define critical thinking. Keep in mind that no single definition exists; it is the responsibility of individual professors and their programs to determine what they mean by "critical thinking."

Critical thinking has sometimes been viewed as identical to logical reasoning. However, logical reasoning is increasingly seen as necessary, but not sufficient, for critical thinking (e.g., Gainen & Locatelli, 1995, p.84).

In a comprehensive monograph on critical thinking in higher education, Kurfiss (1988, p.2) defined critical thinking as follows:

> *Critical thinking is...an investigation whose purpose is to explore a situation, phenomenon, question, or problem to arrive at a hypothesis or conclusion about it that integrates all available information and that can therefore be convincingly justified.*

Kurfiss' definition assumes that the goal of critical thinking is for students to adequately address *open-ended* or *unstructured* problems—those having no single "correct" solution.[36] This view is consistent with the goals of most business educators, who would like students to be able to address open-ended problems such as the following:

- Create a marketing plan;
- Interpret financial statement ratios;
- Recommend an organizational structure;
- Identify and address underlying causes for business problems;
- Analyze a company's strengths, weaknesses, opportunities, and threats;
- Recommend improvements to an information system; and
- Anticipate employee behavioral effects of a planned business decision
.

Although Kurfiss' general definition is helpful, it does not provide a very complete description of critical thinking skills. A more explicit description would provide a better basis for assessment and also improve communication for teaching and learning.

Sometimes professional organizations offer guidance about the types of skills they would like students to exhibit upon graduation. For example, the American Institute of Certified Public Accountants (AICPA) has published a suggested list of 20 competencies in its *Core Competency Framework for Entry Into the Accounting Profession*. Column (a) of Figure 1 presents elements for one of the competencies, called "Problem Solving and Decision

Figure 1
Examples of Critical Thinking Skills

(a)	(b)	(c)
AICPA (1999; 2002)	**Braun (2004)**	**Baril, *et al.* (1998)**
Personal Competency: Problem Solving and Decision Making	**Critical Thinking Skills Applied to a Case Problem**	**Public Accountants' Definition of Critical Thinking**
• Lists information and evidence that is relevant for a problem • Identifies uncertainties about the interpretation or significance of information and evidence • Makes valid and reliable evaluations of information, including the significance of evidence or facts for problem definition and solution • Considers unconventional approaches and solutions to problems • Analyzes the impact, pros, and cons of potential solutions or actions • Analyzes the quality of information and evidence, including validity, reliability, and significance • Synthesizes novel or original definitions of problems and solutions as circumstances dictate • Uses experience and comparison in forming opinions • Seeks consensus where appropriate • Reasons carefully and thinks effectively in abstract terms or generalizations	• Know where and how to get information • Determine facts and evidence • Recognize central thesis or arguments • Assemble information into one's own words • Assess relevance, truth, validity, and strength of information • Recognize stated and implied assumptions • Identify cause and effect relationships • Evaluate information from multiple perspectives • Draw conclusions and evaluate the effectiveness of justification of the conclusion • Evaluate information from multiple perspectives • Create multiple options • Determine the criteria to use to evaluate options • Assess strengths and weaknesses of options	**Cognitive Attributes and Characteristics:** • Recognizes problem areas • Recognizes when additional information is needed • Fits details into the overall environment; sees the "big picture" • Transfers knowledge from one situation to another • Anticipates, thinks ahead, plans **Non-Cognitive Attitudes and Behaviors** • Exhibits initiative • Exhibits curiosity • Exhibits confidence • Communicates clearly and articulately **Other:** • Displays creativity • Accepts ambiguity • Recognizes when there is more than one acceptable solution • Makes qualitative judgments • Displays rapid thought process

Figure 1 (continued)
Examples of Critical Thinking Skills

(a)	(b)	(c)
AICPA (1999; 2002)	**Braun (2004)**	**Baril, *et al.* (1998)**
Personal Competency: Problem Solving and Decision Making	**Critical Thinking Skills Applied to a Case Problem**	**Public Accountants' Definition of Critical Thinking**
• Knows when to follow directions, question plans or seek help • Strategically considers contingencies and future developments • Adapts to new contexts and promotes constructive change		• Displays healthy skepticism; asks "why?" or "why not?" • Challenges the status quo • Determines the extent of what is reasonable; defines the limits of acceptability • Recognizes personal limitations • Exposed to diverse cultures, knowledge and backgrounds • Recognizes presence of biases

Making." These elements provide a fairly complete list of desirable critical thinking skills that is consistent with Kurfiss' definition.

Academic researchers who focus on how to teach critical thinking skills often provide lists of desired skills. For example, column (b) of Figure 1 presents a list of critical thinking skills for business case analysis derived from Braun (2004, Table 1). This list is also consistent with Kurfiss' definition of critical thinking.

Another approach for identifying skills is to gather information from employers or other interested stakeholders. Column (c) of Figure 1 presents a list of critical thinking skills based on a content analysis of interviews with accounting professionals at seven public accounting firms. This list of skills is broader than the other two sets in Figure 1; it includes skills such as "exhibits initiative" and "displays creativity" that are often viewed as separate from critical thinking. Within academia, disagreements exist about whether attributes such as creativity are part of critical thinking (e.g., Bailin, 1993).

The differences among sets of critical thinking skills shown in Figure 1 highlight the importance of defining the meaning of critical thinking before attempting to assess it. At the program level, this could be engaging in conversations among faculty members to reach a consensus, taking into account the program's mission. At the individual course level, professors should clarify the critical thinking skills they believe are important in the course, taking into account the skills called for in the program. The process of reflecting explicitly upon desired skills often leads to improved course design and communication with students, which in turn can improve student performance.

Learning Goals

Once critical thinking skills are adequately described, the next step is to establish goals for student critical thinking performance, often called *learning goals* or *desired learning outcomes*.[37] Learning goals fall into two general categories: *norm-referenced* and *criterion-referenced*. Norm-referenced goals refer to the ranking of student performance relative to other students. For example, a program may establish a learning goal that students in the program rank in the 75th percentile on a nationally-normed critical thinking test. Criterion-referenced goals refer to whether students have achieved specific skills, with performance measured against preset standards. For example, a program may establish a learning goal that at least 90% of students in the program are able to adequately perform a specific list of critical thinking skills.

There is currently an increased trend in higher education to establish criterion-referenced learning goals, which is consistent with a shift toward competency-based education. Programs and individual professors are increasing their focus on the skills and abilities required of students upon graduation or upon completing a course. This approach is particularly relevant in the area of critical thinking; the goal is for students to be able to think critically. Thus, the remainder of this section focuses on criterion-referenced learning goals for critical thinking. The process of establishing criterion-referenced learning goals naturally begins with a description of critical thinking (as discussed in the preceding section). However, educators do not expect the same critical thinking performance from students in two-year, four-year, master's, and doctoral programs. Nor do they expect students at the beginning of a program to perform as well as students at the end of a program. How should they decide which skills are appropriate as learning goals for an individual program or course?

Using their considerable experience with students and their understanding of a program's mission, faculty could discuss potential objectives and use their collective judgment to decide criteria for acceptable and unacceptable performance for each of the critical thinking skills. Some skills, such as "lists information and evidence that is relevant for a problem" might be considered essential for students in all programs, whereas a skill

such as "strategically considers contingencies and future developments" might be considered essential only for students in a graduate-level program. Assessments of student performance could be used to assist this process (discussed further later in this chapter).

A model of cognitive development could be used more formally to assist faculty in identifying appropriate critical thinking skills for students in a given program or course. A range of cognitive development models exist. However, Wolcott *et al.* (2002, p.92) point out that the various models have the following themes in common:

- Critical thinking skills can be arrayed cognitively from less complex to more complex;
- Students must develop less complex skills before they can develop more complex skills;
- Most college students operate at cognitive levels that are too low for adequate critical thinking performance;
- Critical thinking skills develop slowly (if they do develop); and
- Cross-curricular educational efforts, and educational efforts over time, are needed to give students sufficient time and practice for development of critical thinking skills.

In other words, a model of cognitive development can be used to formally sort critical thinking skills from less to more complex. Then a learning goal can be established based on the level of skills students are expected to achieve in a particular program. Learning goals can also be established for intermediate points in a program and for individual courses. This approach for setting learning goals is particularly useful for assessment purposes. The bottom corner of the assessment triangle shown in Figure 2 (adapted from Pellegrino *et al.*, 2001, pp.44-51, 296) indicates that assessments should be grounded in a model of cognition and learning, which should explicitly tie assessments to an understanding of how student skills are expected to develop. This approach is particularly useful because it facilitates "closing the assessment loop"—using assessment results to inform teaching and learning. Because models of cognitive development provide guidance about how critical thinking skills are likely to develop, they provide an appropriate basis for conducting assessments of these skills.[38]

To facilitate setting learning goals according to level of complexity, the AICPA has organized the elements for each of its 20 competencies into levels 1, 2, 3, and 4, based on the stages of cognitive development defined by the reflective judgment model.[39] These levels allow faculty to quickly identify which skills are necessary for development of more complex skills, and it provides a cohesive way to establish learning goals for a wide range of competencies.

The process of sorting the competency elements prompted a revision

Figure 2
Assessment Triangle

```
┌─────────────────────────────┐          ┌─────────────────────────────┐
│        Observation          │          │       Interpretation        │
│     A set of beliefs about  │          │      An  interpretation     │
│  the kinds of observations  │          │   process for making sense  │
│  that will provide evidence of│        │   of assessment evidence    │
│  students' competencies     │          │                             │
└─────────────────────────────┘          └─────────────────────────────┘
              \                                      /
               \                                    /
                \                                  /
                 \                                /
                  \                              /
                   \                            /
                    ┌─────────────────────────────┐
                    │         Cognition           │
                    │      A model of student     │
                    │   cognition and learning in │
                    │        the domain           │
                    └─────────────────────────────┘
```
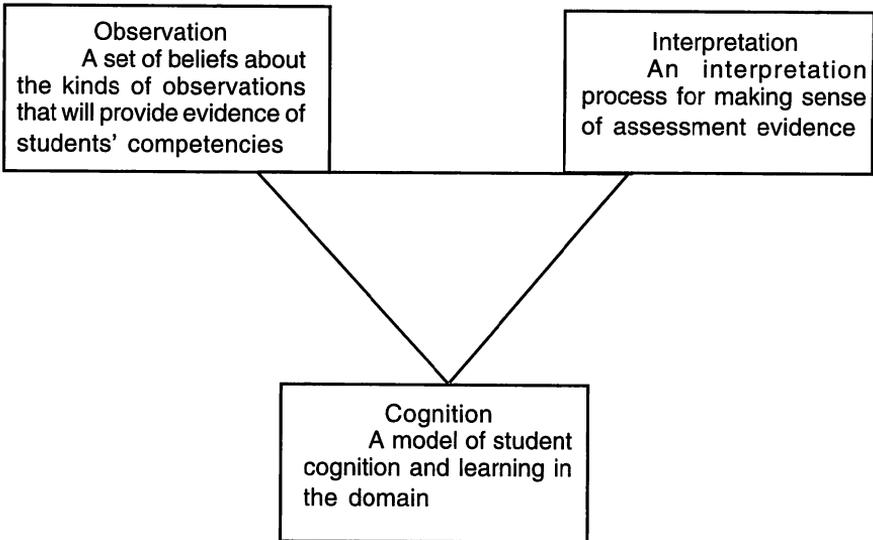
Adapted from Pellegrino *et al.* (2001, pp.44-51, 296)

in the AICPA's list of elements for many of its competencies. For example, panels A and B of Figure 3 show the original and revised list of elements for the Problem Solving and Decision Making competency.[40] Notice that the original list did not include any skills at level 1. Most similar lists of skills fail to include descriptions at the lowest levels, causing faculty and students to pay insufficient attention to development of the least complex skills, which lay the essential foundation for higher-level skills. Also notice that descriptions for several of the skills were revised to clarify the specific requirements. Providing more precise descriptions improves the ability to assess the skills and also improves communication between students and faculty about desired skills.

For the critical thinking skills listed in Panel B of Figure 3, different programs would be likely to choose different levels as their learning goals. For example, a two-year program might choose Level 1 as the learning goal, while a four-year program might choose Level 2, a master's program Level 3, and a doctoral program Level 4. The specific levels that would be appropriate depend on both the characteristics of students in the program and the program's mission.[41]

Assessment Methods

Just as there is no single way to define critical thinking, there is no single way to assess it. All assessment methods have strengths and

Figure 3
Levels of Complexity for Skills in AICPA Core Competency: Problem Solving and Decision Making
Panel A: Original List of Elements (AICPA, 1999)

← Less Complex		More Complex →	
Level 1	Level 2	Level 3	Level 4
	• Verifies information for problem definition and solution. • Considers unconventional approaches and solutions to problems. • Makes valid and reliable evaluations of information. • Evaluates the significance of evidence or facts	• Synthesizes novel or original definitions of problems and solutions as circumstances dictate. • Uses experience and comparison in forming opinions. • Proposes and evaluates alternative solutions. • Seeks consensus where appropriate. • Reasons carefully and thinks effectively in abstract terms or generalizations. • Analyzes the impact of potential actions. • Knows when to follow directions, question plans, or seek help.	• Considers contingencies and future developments • Adapts to new contexts and promotes constructive change.

weaknesses, and judgment is required in choosing one or more methods that are most appropriate for a particular program or course. The Assessment Triangle introduced in the last section (Figure 2) is a useful tool for thinking about the design and interpretation of critical thinking assessment. The assessment begins with a model of cognition, which helps in measuring student performance and also suggests ways to improve teaching and learning.

Figure 3 (continued)
Panel B: Revised List of Elements (AICPA 2002)

← Less Complex		More Complex →	
Level 1	Level 2	Level 3	Level 4
• Lists information and evidence that is relevant for a problem. • Identifies uncertainties about the interpretation or significance of information and evidence.	• Makes valid and reliable evaluations of information, including the significance of evidence or facts for problem definition and solution. • Considers unconventional approaches and solutions to problems. • Analyzes the impact, pros, and cons of potential solutions or actions. • Analyzes the quality of information and evidence, including validity, reliability, and significance.	• Synthesizes novel or original definitions of problems and solutions as circumstances dictate. • Uses experience and comparison in forming opinions. • Seeks consensus where appropriate. • Reasons carefully and thinks effectively in abstract terms or generalizations. • Knows when to follow directions, question plans, or seek help.	• Strategically considers contingencies and future developments. • Adapts to new contexts and promotes constructive change.

The next corner focuses on the types of assessment observations that will provide reasonable data. Some assessment methods provide better evidence about critical thinking than other methods. The final corner focuses on interpretation, which is required for any type of research. Interpretation might include formal quantitative or statistical analysis and/or qualitative evaluations of the assessment setting, design, and results. Several key factors influence the choice of assessment methods:[42]

- *Validity*: The degree to which the assessment method measures what it is supposed to measure—in this case, student critical thinking. Validity depends not only on the how well the assessment is constructed, but also on how well it matches the program/course critical thinking definition and learning goal.

138

- *Reliability*: The degree to which the assessment method minimizes measurement errors so that variance in scores can be attributed to actual differences in student critical thinking.
- *Intended Use*: The ability of the assessment method to provide information that is useful for the intended purpose and is understandable to the planned audience.

Table 1 provides examples as well as major strengths and weaknesses of four major types of methods that may be used to assess critical thinking skills: objective tests, performances, surveys, and proxy indicators. Many experts argue that critical thinking skills are multi-faceted and that performances—such as essays, cases, and simulations—are the best way to assess them (e.g., Gainen & Locatelli 1995, p.85).[43] In addition, performance assessments may provide the best fit with the assessment triangle; they can be designed in conjunction with a cognitive development model so that assessment results directly guide improvements in teaching and learning. Many programs prefer using multiple assessment methods for outcomes such as critical thinking. A portfolio of assessments is likely to provide the best information, and the results of one method can be used to confirm or better understand the results of another method. Nevertheless, due to space limitations the rest of this section will focus only on performance assessments that are locally developed for a course or program.

In a performance assessment, students are given a task that requires them to demonstrate critical thinking in a product such as an essay, written case analysis, or oral presentation.[44] Their performances are then assessed by faculty or other competent raters with the aid of a set of criteria, called a rubric. There are two general types of rubrics: holistic and analytical. The discussion that follows provides examples of each. Individual programs and instructors may wish to adopt an existing rubric or create their own.[45] The choice of rubric depends on the same factors that affect the choice of assessment method, including validity, reliability, and intended use.

Figure 4 presents a holistic rubric for critical thinking, created by Peter and Noreen Facione, which describes four possible levels of performance. A holistic rubric assigns a level of performance based on evaluation of a set of criteria, taken as a whole. Thus, each performance receives only a single score. This type of rubric is fairly easy to use and helps raters focus on each student's global performance.

Table 2 presents an analytical rubric for critical thinking created by a committee chaired by Thomas Calderon at the University of Akron for use on accounting case assignments. An analytical rubric assigns levels of performance on two or more criteria. Calderon's rubric provides for ratings on 11 critical thinking traits. This type of rubric takes more time to complete, but it provides more details about each performance. Raters are allowed to give

Table 1
Overview of Critical Thinking Assessment Methods

Type	Examples	Major Pros and Cons
Objective Tests	Watson Glaser Critical Thinking Appraisal California Critical Thinking Dispositions Inventory California Critical Thinking Skills Test Cornell Critical Thinking Test College Outcomes Measures Program - Objective Test Locally-developed tests	• Efficient and inexpensive way to collect data. • Can apply well-established practices for evaluating reliability and validity. • May allow comparison with external norms • Often focus on logical reasoning, which is insufficient for measuring desired critical thinking skills • Labor intensive to develop a valid and reliable test. • Participation rates and/or motivation may be low if ungraded.
Performances	ETS Tasks in Critical Thinking Reflective Judgment Interview Ennis-Weir Critical Thinking Essay Test Coursework assignments: Essay Case Simulation Project Oral presentation	• Efficient way to collect data if combined with regular coursework. • More direct and thorough assessment of complex, integrated critical thinking skills. • More authentic; can simulate "real-life" applications. • Promotes integration of teaching, learning, and assessment when combined with coursework. • Labor-intensive to develop and use • May lack validity and reliability (depending on design). • Participation rates and/or motivation may be low if ungraded.
Surveys	Student exit survey Alumni survey Employer survey Faculty survey Focus groups	• Efficient and inexpensive way to collect data. • Provide useful information about perceptions and experiences. • Can apply well-established practices for evaluating reliability and validity. • Do not directly measure critical thinking skills. • Participation rates and motivation may be low.
Proxy Indicators	Course grades Analysis of course syllabi, assignments, exams Post-graduate employment statistics Graduate school admissions rates Internship participation rates	• Efficient and inexpensive way to collect data. • Do not directly measure critical thinking skills. • Results may be difficult to interpret.
Adapted from Wolcott (2002)		

Figure 4
Hooistic Critical Thinking Scoring Rubric
Facione and Facione

4. Consistently does all or almost all of the following:
    Accurately interprets evidence, statements, graphics, questions, etc.
    Identifies the salient arguments (reasons and claims) pro and con.
    Throughtfully analyzes and evaluates major alternative points of view.
    Draws warranted, judicious, non-fallacious conclusions.
    Justifies key results and procedures, explains assumptions and reasons.
    Fair-mindedly follows where evidence and reasons lead.

3. Does most or many of the following:
    Accurately interprets evidence, statements, graphics, questions, etc.
    Identifies relevant arguments (reasons and claims) pro and con.
    Offers analyses and evaluations of obvious alternative points of view.
    Justifies some results or procedures, explains reasons.
    Fair-mindedly follows where evidence and reasons lead.

2. Does most or many of the following:
    Misinterprets evidence, statements, graphics, questions, etc.
    Fails to identify strong, relevant counter-arguments.
    Ignores or superficially evaluates obvious alternative points of view.
    Draws unwarranted of fallacious conclusions.
    Justifies few results or procedures, seldom explains reasons.
    Regardles of the evidence or reasons, maintains or defends views based
        on self-interenst or preconceptions.

1. Consistently does all or almost all of the following:
    Offers biased interpretations of evidence, statements, graphics,
        questions, information, or the points of view of others.
    Fails to identify or hastily dismisses strong, relevant counter-arguments.
    Ignores or superficially evaluates obvious alternative points of view.
    Arugues using fallacious or irrelevant reasons, and unwarranted claims.
    Does not justify results or procedures, nore explain reasons.
    Regardless of the evidence or reasons, maintains or defends views
        based on self-interest or preconceptions.
    Exhibits close-mindedness or hostility or reason.

students different scores on different criteria, providing faculty and students with better diagnostic information for teaching and learning.

Table 3 presents a combined holistic and analytical rubric, created by Cindy Lynch and me, that is based on a teaching and learning model called Steps for Better Thinking. The rubric supports ratings for eight individual criteria (rows A through H), and then an overall assessment associated with five

Table 2
Rubric for Accounting Case Analysis Traits

Traits	Very Good 3	Good/Satisfactory 2	Poor/Unsatisfactory 1
Identify issues	Addresses all major issues and most of the minor ones	Addresses most of the major issues but omits a few of the minor ones	Issues are not addressed, or they are addressed inappropriately
Identify accounting issues and concepts	Addresses all the major accounting issues/concepts and most of the minor ones	Addresses most of the major accounting issues/concepts appropriately but omits or inappropriately identifies a few of the minor ones	Accounting issues/concepts are not identified or identified inappropriately
Identify alternative options	Identifies a set of feasible options that demonstrates creativity and the ability to integrate knowledge	Identifies and discusses a set of feasible alternatives	Does not identify alternatives
Identify criteria	Presents and discusses very thoroughly a coherent set of criteria	Presents and discusses at least briefly a coherent set of criteria	Does not present a coherent set of criteria that are tied to the context of the case

Table 2 (continued)
Rubric for Accounting Case Analysis Traits

Traits	Very Good 3	Good/Satisfactory 2	Poor/Unsatisfactory 1
Quantitative evaluation	Evaluation is effective, consistent with the criteria, and facts are used correctly	Evaluation is adequate, relatively consistent with the criteria, and facts are generally used correctly	Evaluation is poor, not consistent with the criteria or does not use facts correctly
Qualitative evaluation	Evaluation is effective, demonstrates creativity, coherent, consistent with the criteria, and facts are used correctly	Evaluation is adequate, coherent, relatively consistent with the criteria, and facts are generally used correctly	Evaluation is poor, not consistent with the criteria, incoherent or does not use facts correctly
Select an option that is based on the evaluation	A very strong and visible link between analysis and the option selected	Relatively adequate link between the analysis and the option selected	No link between evaluation and option selected
Present and organize work in a logical manner	Analysis is very well organized and flow is effective	Analysis is reasonably well organized and flow is adequate	Analysis is poorly organized and does not flow effectively
Use new ideas and analysis methods not included in the case	Effectively integrates ideas and analytical methods from other sources not directly presented in the case	Integrates ideas and analytical methods from other sources not directly presented in the case reasonably well	Absence of any new ideas and creative analytical methods
Use ethical and profession-ally responsible documenta-tion and propose ethical and responsible solutions	Sources are appropriate, well documented and effectively cited. Where appropriate the decision is ethically sound, credible, and not frivolous.	Sources are fair, reasonably well documented and cited. Where appropriate the decision is ethically sound, credible, and not frivolous.	Sources are inappropriate, not well documented and poorly cited. Decisions are not ethically sound, credible, or are frivolous.
Source: Thomas Calderon, University of Akron, Akron, OH 44325-4802; tcalderon@uakron.edu			

levels of cognitive development (performance patterns 0 through 4).[46] This rubric is highly detailed, which makes it more difficult to learn to use but also allows for improved assessment reliability. The simple version of this rubric in Table 4 is a better tool for classroom use, because it is quicker to use and provides better communication with students. However, this version is not as reliable, because it provides less complete descriptions of the assessment criteria.

The rubrics presented in Figure 4 and Tables 2-4 are designed for generic use across multiple performance tasks. Rubrics may also be created for use with individual tasks. For example, Table 5 presents a version of the rubric from Table 4 that has been modified for a textbook mini-case assignment called *Hudziak Industries* (Eldenburg & Wolcott, 2005). Customized rubrics are easier to use than generic ones, but they might reduce students' abilities to recognize that the same set of critical thinking skills are called for across different tasks.

Implementation Issues

Implementing a performance assessment involves the following steps: defining desired skills and learning goals, designing and implementing an assessment task, choosing or designing an assessment rubric, rating and summarizing student performances, interpreting assessment results in

Table 3
Steps for Better Thinking Rubric

	Less Complex Performance Patterns		More Complex Performance Patterns		
	"Confused Fact Finder" Performance Pattern 0-How performance might appear when Step 1, 2, 3, and 4 skills are weak	"Biased Jumper" Performance Pattern 1ÑHow performance might appear when Step 1 skills are adequate, but Step 2, 3, and 4 skills are weak	"Perpetual Analyzer" Performance Pattern 2-How performance might appear when Step 1 and 2 skills are adequate, but Step 3 and 4 skills are weak	"Pragmatic Performer" Performance Pattern 3-How performance might appear when Step 1, 2, and 3 skills are adequate, but Step 4 skills are weak	"Strategic Re-Visioner" Performance Pattern 4-How performance might appear when one has strong Step 1, 2, 3, and 4 skills
Steps for Better Thinking **SKILLS**					
Step 1: **IDENTIFY** A-Identify and use relevant information B-Articulate uncertainties	A0-Uses very limited information; primarily "facts," definitions, or expert opinions B0-Either denies uncertainty OR attributes uncertainty to temporary lack of information or to own lack of knowledge	A1-Uses limited information, primarily evidence and information supporting own conclusion* B1-Identifies at least one reason for significant and enduring uncertainty*	A2-Uses a range of carefully evaluated, relevant information B2-Articulates complexities related to uncertainties and the relationships among different sources of uncertainty	A3-Uses a range of carefully evaluated, relevant information, including alternative criteria for judging among solutions B3-Exhibits complex awareness of relative importance of different sources of uncertainties	A4-Same as A3 PLUS includes viable strategies for GENERATING new information to address limitations B4-Exhibits complex awareness of ways to minimize uncertainties in coherent, on-going process of inquiry

Table 3 (continued)
Steps for Better Thinking Rubric

Step 2: EXPLORE	C0-Portrays	C1-Acknowledges	C2-Interprets	C3-Evaluates	C4-Same as C3
C-Integrate multiple perspectives and clarify assumptions D-Qualitatively interpret information and create a meaningful organization	perspectives and information dichotomously, e.g., right/wrong, good/bad, smart/stupid D0-Does not acknowledge interpretation of information; uses contradictory or illogical arguments; lacks organization	more than one potential solution, approach, or viewpoint; does not acknowledge own assumptions or biases D1-Interprets information superficially as either supporting or not supporting a point of view; ignores relevant information that disagrees with own position; fails to sufficiently break down the problem	information from multiple viewpoints; identifies and evaluates assumptions; attempts to control own biases* D2-Objectively analyzes quality of information; organizes information and concepts into viable framework for exploring realistic complexities of the problem*	information using general principles that allow comparisons across viewpoints; adequately justifies assumptions D3-Focuses analyses on the most important information based on reasonable assumptions about relative importance; organizes information using criteria that apply across different viewpoints and allow for qualitative comparisons	PLUS argues convincingly using a complex, coherent discussion of own perspective, including strengths and limitations D4-Same as D3 PLUS systematically reinterprets evidence as new information is generated over time OR describes process that could be used to systematically reinterpret evidence
Step 3: PRIORITIZE E-Use guidelines or principles to judge objectively across the various options F-Implement and communicate conclusions for the setting and audience	E0-Fails to reason logically from evidence to conclusions; relies primary on unexamined prior beliefs, clichés, or an expert opinion F0-Creates illogical implementation plan; uses poor or inconsistent communication; does not appear to recognize existence of an audience	E1-Provides little evaluation of alternatives; offers partially reasoned conclusions; uses superficially understood evidence and information in support of beliefs F1-Fails to adequately address alternative viewpoints in implementation plans and communications; provides insufficient information or motivation for audience to adequately understand alternatives and complexity	E2-Uses evidence to reason logically within a given perspective, but unable to establish criteria that apply across alternatives to reach a well-founded conclusion OR unable to reach a conclusion in light of reasonable alternatives and/or uncertainties F2-Establishes overly complicated implementation plans OR delays implementation process in search of additional information; provides audience with too much information (unable to adequately prioritize)	E3-Uses well-founded, overarching guidelines or principles to objectively compare and choose among alternative solutions; provides reasonable and substantive justification for assumptions and choices in light of other options* F3-Focuses on pragmatic issues in implementation plans; provides appropriate information and motivation, prioritized for the setting and audience*	E4-Articulates how a systematic process of critical inquiry was used to build solution; identifies how analysis and criteria can be refined, leading to better solutions or greater confidence over time F4-Implementation plans address current as well as long-term issues; provides appropriate information and motivation, prioritized for the setting and audience, to engage others over time

Table 3 (continued)
Steps for Better Thinking Rubric

Step 4: ENVISION G-Acknowledge and monitor solution limitations through next steps H-Overall approach to the problem	G0-Does not acknowledge significant limitations beyond temporary uncertainty; next steps articulated as finding the "right" answer (often by experts) H0-Proceeds as if goal is to find the single, "correct" answer	G1-Acknowledges at least one limitation or reason for significant and enduring uncertainty; if prompted, next steps generally address gathering more information H1-Proceeds as if goal is to stack up evidence and information to support own conclusion	G2-Articulates connections among underlying contributors to limitations; articulates next steps as gathering more information and looking at problem more complexly and/or thoroughly H2-Proceeds as if goal is to establish an unbiased, balanced view of evidence and information from different points of view	G3-Adequately describes relative importance of solution limitations when compared to other viable options; next steps pragmatic with focus on efficiently GATHERING more information to address significant limitations over time H3-Proceeds as if goal is to come to a well-founded conclusion based on objective consideration of priorities across viable alternatives	G4-Identifies limitations as in G3; as next steps, suggests for strategically GENERATING new information to aid in addressing significant limitations over time* H4-Proceeds as if goal is to strategically construct knowledge, to move toward better conclusions or greater confidence in conclusions as the problem is addressed over time*

2003, Susan K. Wolcott. Permission is granted to reproduce this information for noncommercial purposes. Please cite this source: Wolcott, S. K. (October 29, 2003). Steps for Better Thinking Rubric [online]. Available: http://www.WolcottLynch.com. Based in part on information from Reflective Judgment Scoring Manual With Examples (1985/1996) by K. S. Kitchener and P. M. King. Grounded in dynamic skill theory (Fischer & Bidell, 1998).

Table 4
Steps for Better Thinking Competency Rubric

		Performance Pattern 0 "Confused Fact Finder"	Performance Pattern 1 "Biased Jumper"	Performance Pattern 2 "Perpetual Analyzer"	Performance Pattern 3 "Pragmatic Performer"	Performance Pattern 4 "Strategic Revisioner"
Step 1 Skills	Identify relevant information Circle ALL that apply	Identifies facts, definitions, and/or experts' opinions.	Identifies information[1] that is relevant[1] to the problem	Explores a wide range of relevant information[1]	Focuses on the most important relevant information[1]	Develops viable strategies for generating important relevant information[1] over time
	Recognize and address uncertainties[2] Circle ALL that apply	Identifies at least one reason for temporary uncertainty[2]	Identifies at least one reason for significant and permanent uncertainty[2]	Addresses significant and permanent uncertainties[2] when interpreting information	Identifies and discusses the significance of the most important uncertainties[2]	Develops viable strategies for minimizing important uncertainties[2] over time
Step 2 Skills	Integrate multiple perspectives[3] and clarify assumptions[4] Circle ALL that apply		Acknowledges more than one potential solution, approach, or viewpoint	Analyzes information from multiple perspectives,[3] including assumptions[4] alternative objectives	Provides reasonable and substantive justification for assumptions4 used in analysis	Argues convincingly using a complex, coherent discussion of own perspective; Articulates strengths and weaknesses of position
	Interpret and organize information Circle ALL that apply		Uses evidence logically to support a point of view; Correctly applies concepts/theories/techniques	Qualitatively interprets information and develops meaningful categories for analysis	Preserves problem complexity, but emphasizes the most important and/or most relevant and reliable information	Systematically re-interprets information as circumstances change or new information becomes available

146

Table 4 (continued)
Steps for Better Thinking Competency Rubric

Step 3 Skills	Use guidelines or principles to judge objectively across options Circle ALL that apply		Avoids reaching a biased conclusion	Maintains objectivity while establishing reasonable priorities for reaching a well-founded conclusion	Uses a systematic process of critical inquiry to build a solution; Articulates how problem solving approach and criteria can be refined, leading to better solutions or greater confidence over time
	Communicate and implement conclusions Circle ALL that apply			Appropriately tailors communication or implementation plans to the setting and audience	Provides appropriate information to motivate and engage others in long-term strategies
	Address solution limitations Circle ALL that apply			Focuses on most efficient ways to address limitations or to gather additional information	Articulates solution limitations as a natural part of addressing open-ended problems
Step 4 Skills	Engage in continuous improvement Circle ALL that apply				Identifies uncertainties and limitations as opportunities for continuous improvement; Engages in lifelong learning
Overall Approach to the Problem Circle ONLY ONE	Proceeds as if goal is to find the single, "correct" answer	Proceeds as if goal is to stack up evidence and information to support own conclusion	Proceeds as if goal is to establish an unbiased, balanced view of evidence and information from different points of view	Proceeds as if goal is to come to a well-founded conclusion based on objective consideration of priorities across viable alternatives	Proceeds as if goal is to strategically construct knowledge, to move toward better conclusions and greater confidence in conclusions as the problem is addressed over time

[1] Information can take many forms, including facts, descriptions, definitions, arguments, opinions, ideas, claims, theories, concepts, observations, research findings, values, perceptions, beliefs, influences, effects, and so on. Information can be obtained in many ways such as reading, seeing, hearing, touching, feeling, experiencing, interacting, thinking, etc.

[2] Uncertainties can relate to many aspects of the problem, including the problem definition, availability of solution alternatives, quality and interpretation of information, effects of alternatives, priorities and values of the decision maker and others, and so on.

[3] Perspectives can relate to any type of grouping that is meaningful to the problem, such as categories of people, cultures, societies, roles, races, genders, hierarchies, theories, concepts, ideas, beliefs, attitudes, physical locations, time, disciplines, values, emotions, and so on.

[4] Assumptions are hypotheses, suppositions, conjectures, assertions, presumptions, beliefs, or premises that are taken for granted or that lie behind an argument. Assumptions are made because of uncertainties; the "truth" cannot be known or proven. Some assumptions are better than others. Better assumptions are more reasonable, logical, comprehensive, plausible, likely, rational, impartial, objective, justified, credible, and/or believable.

[©] 2003, Susan K. Wolcott. Permission is granted to reproduce this information for noncommercial purposes. Please cite this source: Wolcott, S. K. (October 29, 2003). Steps for Better Thinking Competency Rubric [online]. Available: http://www.WolcottLynch.com. Based in part on information from Reflective Judgment Scoring Manual With Examples (1985/1996) by K. S. Kitchener and P. M. King. Grounded in dynamic skill theory (Fischer & Bidell, 1998).

conjunction with the results of other relevant assessments, taking actions based on assessment results, and refining the assessment process over time. Below are key issues to consider beyond those addressed earlier in this chapter.

Designing and Implementing an Assessment Task: An assessment task (such as a case or other assignment) must be chosen that will encourage students to demonstrate the critical thinking skills defined for the program or course. Thus, a task that is appropriate in one setting may be inappropriate

Table 5: Steps for Better Thinking Competency Rubric: Hudziak Industries*

		Performance Pattern 0 "Confused Fact Finder"	Performance Pattern 1 "Biased Jumper	Performance Pattern 2 "Perpetual Analyzer"	Performance Pattern 3 "Pragmatic Performer"	Performance Pattern 4 "Strategic Revisioner"
Overall Approach to the Problem Circle ONLY ONE		Proceeds as if goal is to find the single, "correct" answer	Proceeds as if goal is to stack up evidence and information to support own conclusion	Proceeds as if goal is to establish an unbiased, balanced view of evidence and information from different points of view	Proceeds as if goal is to come to a well-founded conclusion based on objective consideration of priorities across viable alternatives	Proceeds as if goal is to strategically construct knowledge, to move toward better conclusions or greater confidence in conclusions as the problem is addressed over time
Step 1 Skills	Identify relevant information Circle ALL that apply	Appropriately identifies the most profitable option for the company and for each profit center	Identifies at least some information[1] beyond calculations that is relevant to the raw material and manager conflict decisions	Explores a wide range of relevant information,[1] including information not explicitly presented in the problem	Focuses on the most important relevant information[1] for making the raw material and manager conflict decisions	Develops viable strategies for generating important relevant information[1] over time
	Recognize and address uncertainties[2] Circle ALL that apply	Identifies at least one temporary uncertainty[2] about the raw material or manager conflict decisions	Identifies at least one significant uncertainty[2] about the raw material and manager conflict decisions	Addresses uncertainties[2] when discussing options for the raw material and manager conflict decisions	Identifies and discusses the significance of the most important uncertainties[2]	Develops viable strategies for minimizing/addressing important uncertainties[2] over time
Step 2 Skills	Integrate multiple perspectives[3] and clarify assumptions[4] Circle ALL that apply		Acknowledges more than one potential solution to the raw material and manager conflict decisions	Addresses multiple perspectives[3] and assumptions[4] when discussing options for the raw material and manager conflict decisions	Provides reasonable and substantive justification for assumptions[4] used in analysis	Argues convincingly using a complex, coherent discussion of own perspective; Articulates strengths and weaknesses of position
	Interpret and organize information Circle ALL that apply		Uses evidence logically to support recommendations; Correctly applies concepts/theories/ techniques	Addresses the quality of information; Organizes the memo meaningfully	Preserves complexity and objectivity, but emphasizes the most important and/or most relevant and reliable information	Explains how information might be reinterpreted as circumstances change or new information becomes available

		Performance Pattern 0 "Confused Fact Finder"	Performance Pattern 1 "Biased Jumper	Performance Pattern 2 "Perpetual Analyzer"	Performance Pattern 3 "Pragmatic Performer"	Performance Pattern 4 "Strategic Revisioner"
Overall Approach to the Problem Circle ONLY ONE		Proceeds as if goal is to find the single, "correct" answer	Proceeds as if goal is to stack up evidence and information to support own conclusion	Proceeds as if goal is to establish an unbiased, balanced view of evidence and information from different points of view	Proceeds as if goal is to come to a well-founded conclusion based on objective consideration of priorities across viable alternatives	Proceeds as if goal is to strategically construct knowledge, to move toward better conclusions or greater confidence in conclusions as the problem is addressed over time
Step 3 Skills	Use guidelines or principles to judge objectively across options Circle ALL that apply			Avoids reaching biased conclusions	Maintains objectivity while establishing reasonable priorities for the raw material and manager conflict decisions	Uses a systematic process of critical inquiry to build a solution; Articulates how problem solving approach and criteria can be refined, leading to better solutions or greater confidence over time
	Communicate and implement conclusions Circle ALL that apply				Appropriately tailors the memo to the setting and audience (non-accountant managers)	Provides appropriate information to motivate and engage others in long-term strategies
	Address solution limitations Circle ALL that apply				Focuses on most efficient ways to address limitations or to gather additional information	Articulates solution limitations as a natural part of addressing open-ended problems
Step 4 Skills	Engage in continuous improvement Circle ALL that apply					Identifies uncertainties and limitations as opportunities for continuous improvement

[The footnotes providing descriptions of information, uncertainties, perspectives, and assumptions, are not copied here.]

*Adapted from Eldenburg, L. E. and S. K. Wolcott. 2005. Instructor's Manual for Cost Management: Measuring, Monitoring, and Motivating Performance. Hoboken, NJ: John Wiley & Sons. Adapted with permission from Wolcott, S. K. (October 29, 2003). *Steps for Better Thinking Competency Rubric.*

in another. Shorter assessment tasks are easier to rate, but longer tasks may provide better information about students' critical thinking skills. Sometimes an assessment task must be modified or replaced to obtain valid, reliable, and useful assessment information. It is often easiest to begin with an assignment that has been used previously. The links among assessment, teaching, and learning are maximized when the task is part of normal coursework. If the course instructor agrees to perform assessments as part of the grading process, less effort is needed for overall assessment. Identical or comparable assessment tasks may be used with different student cohort groups to evaluate the effects of program or course changes over time, or with the same student cohort at different points in the curriculum to evaluate changes in student critical thinking skills during their educational experience.[47]

Rating Student Performances: The assessment rubric should provide sufficient descriptions of student performances, and the individuals who rate performances must be capable of achieving reliable ratings. Two raters should independently assess at least a sample of all responses used for program assessment to allow review of inter-rater reliability—the proportion of ratings upon which the raters agree. Erwin (1991, p.65) recommended that inter-rater reliability should be at least 70%. Poor inter-rater reliability can be caused by an inadequate rubric, rater inconsistencies, or a poorly designed assessment task.

Interpreting Assessment Results: Assessment results require interpretation, similar to the process used when interpreting other types of research results. In the early stages of assessment, the results tend to be mostly descriptive; faculty members learn more about their students' critical thinking skills and how the skills change across the curriculum. They can also use early assessment results as baseline data. As greater experienced is gained with critical thinking assessment, the faculty can begin to formulate and test hypotheses about the impact of curricular or other educational efforts.

Taking Actions Based on Assessment Results: "Closing the assessment loop" is often the area of greatest difficulty in a program assessment plan. Yet, this is probably the most important part of the process. Several of the recommendations made throughout this chapter have been aimed at maximizing the likelihood that assessment results will be used. For example, performing assessments as part of normal coursework involves faculty in the process and helps them take ownership of both the assessment results and the use of those results. For example, a finding that students lack the level 1 skills shown in panel b of Figure 3 may encourage faculty to incorporate those skills more explicitly in their courses. The use of a model of cognition, as shown in Figure 2, further facilitates the improvement of teaching and learning. Providing students with feedback using assessment rubrics can prompt their active participation in developing desired critical thinking skills.

Refining the Assessment Process Over Time: Assessment of critical thinking should evolve over time, because it involves gaining new knowledge. Accordingly, the methods used and insights gained will change over time. In addition, faculty may find that their definitions and learning goals for critical thinking change as they learn more about student skills and how those skills develop. Priorities may also change over time. Initially, it may be most important to quickly establish data about students' critical thinking. As assessment experience is gained, it may become more important to improve the quality of assessments.

Pitfalls to Avoid

One of the biggest pitfalls in assessing critical thinking is a failure to "Just Do It." For example, it may be more important to adopt a set of critical thinking skills (perhaps one created by others) than to spend too much time attempting to establish a "perfect" set. Many program-level assessment efforts have been derailed because faculty failed to reach consensus about desired skills. Similarly, assessment efforts often stagnate because of unrealistic attempts to adopt only the most valid and reliable assessment methods. Making trade-offs is a necessarily part of the assessment process, and it can be easier to make the trade-offs when they are explicitly identified and considered. Thinking of assessment as a long-term process can also be helpful in making short-term trade-offs.

Other major pitfalls include adopting assessment methods that provide little useful information or inadequately communicating assessment results. These pitfalls prevent "closing the assessment loop," because stakeholders either do not care about the results or are not given an opportunity to make use of them. The AACSB currently expects programs to use direct assessment methods—i.e., evidence based on actual work completed by students. Consistent with this expectation, the discussion in this chapter presumes that assessment of critical thinking will be based on direct evaluation of actual work completed by students. Indirect assessment methods (e.g., surveys and focus group discussions) may not produce useful information for understanding the nature of students' critical thinking abilities. Yet these methods are attractive, because they are less costly, easier to use and could produce large amounts of data for sophisticated data analysis.

Conclusions

This chapter provides specific examples of definitions, learning goals, and rubrics for critical thinking. However, the examples shown here should be viewed as suggestions rather than as definitive solutions. Assessment methods should be selected to match the mission and resources available of individual programs and courses, and choices of methods may change over time as greater experience is gained.

Developing and implementing a critical thinking assessment plan can

151

lead to a number of benefits beyond meeting the mandates of accrediting bodies, legislatures, or others. The process of discussing the definition of critical thinking and establishing learning goals can improve consensus among faculty members, which in turn can lead to greater consistency in efforts across courses and improved development of student skills. In addition, explicitly linking assessment to student learning allows for more powerful educational experiences. Student critical thinking skills are most likely to develop when desired skills are clearly articulated, students receive timely and easily understood feedback on their performance, and educational efforts are aimed at levels of cognitive development that are neither too high nor too low. Informative assessment results are also likely to encourage further research efforts to develop better evidence about the effectiveness of teaching and learning methods.

Endnotes

36 Some educators view the term "critical thinking" as referring to the development of a point of view, while "problem solving" is viewed as development of a course of action (e.g., Gainen & Locatelli , 1995, p.85). Other educators view "critical thinking" as applying to open-ended problems, while "problem solving" applies to more well-defined problems having a limited number of potential solutions (e.g., Erwin, 2000, pp.11, 25). Still others view the two terms as interchangeable.

37 Other terms include objectives, intended learning outcomes, desired student outcomes, desired educational accomplishments, intended results, instructional goals, and teaching goals.

38 For overviews of cognitive development models and their relationship to critical thinking skills in higher education, see Kurfiss (1988, pp. 51-68), Pascarella and Terenzini (1991, chapter 4), and Palomba and Banta (1999, pp. 262-263).

39 AICPA levels 1, 2, 3, and 4 correspond to reflective judgment stages 4, 5, 6, and 7. See Pascarella and Terenzini (1991, p. 123) and Hofer and Pintrich (1997, pp. 102-103) for comparisons of the reflective judgment model to other models of cognitive development. See Wolcott & Lynch (2002, chapter 2) and Lynch & Wolcott (2001) for explanations of the link between Steps for Better Thinking, the model used for classification of competency elements by the AICPA, and the stages of cognitive development defined in the reflective judgment model.

40 The elements listed for each competency are not intended to be a complete listing of all possible skills for each competency. However, the AICPA intended the list to provide a representative list of skills.

41 In a meta analysis of 20 years ' research, King and Kitchener (1994, Table 6.6) found that the average first-year college student performed half-way between reflective judgment stages 3 and 4, while the average college senior performed at reflective judgment stage 4 (level 1 in Figure 3). Thus, most undergraduate programs require significant educational effort aimed at the less complex skills shown in Figure 3. Even students in a graduate program

may operate at low levels of cognitive complexity. For example, based on assessments performed in two sections of an introductory financial accounting course, I found that 17% of first-term MBA students lacked level 1 skills.

42 It is not possible in this chapter to provide a complete description of assessment methods or issues to consider. The following resources provide excellent discussions: Erwin (1991 & 2000) and Palomba and Banta (1999).

43 Also see Erwin (2000) for a summary of the specific critical thinking and problem solving skills addressed by common critical thinking assessment tests.

44 Oral presentations might be recorded to facility more reliable rating.

45 An excellent online resource about creating rubrics is available at http:// jonathan.mueller.faculty.noctrl.edu/toolbox/rubrics.htm, authored by Jon Mueller, Professor of Psychology at North Central College. Walvoord & Anderson (1998) provide considerable details about how to establish rating criteria (also called primary trait analysis).

46 For more information about the theoretical underpinnings and instructions for using this rubric, see Wolcott & Lynch (2002, chapter 4). The rubric, instructions, and assessment examples are also available under Educator Resources at http://www.WolcottLynch.com.

47 These approaches also comply with the 9 Principles of Good Practice for Assessing Student Learning published by the American Association for Higher Education (Astin et al., 1996). Those principles state that assessment should focus on learning as it is revealed over time, be ongoing rather than episodic, and involve a wide range of faculty.

References

American Institute of Certified Public Accountants (AICPA). 1999; 2002. *Core competency framework for entry into the accounting profession*. New York, NY: AICPA. Retrieved data, from http://eca.aicpaservices.org/.

Angelo, T.A., & K.P. Cross. (1993). *Classroom assessment techniques: A handbook for college teachers (*2nd Ed.) Jossey-Bass Higher and Adult Education Series. San Francisco, CA: Jossey-Bass.

Astin, A. W., Banta, T. W., Cross, K. P., El-Khawas, E., Ewell, P. T., & Hutchings, P. (1996). *9 Principles of good practice for assessing student learning*. AAHE Assessment Forum. Retrieved data, from http://www.aahe.org.

Bailin, S. (1993, Sep/Oct). Epilogue: Problems in conceptualizing good thinking. *The American Behavioral Scientist, 37* (1), 156-164.

Baril, C. P., Cunningham, B. M., Fordham, D. R., Gardner,R. L. & Wolcott, S. K. (1998). Critical thinking in the public accounting profession: Aptitudes and attitudes. *Journal of Accounting Education, 16* (3/4), 381-406.

Black, H. T., & D. L. Duhon. (2003 Nov/Dec). Evaluating and improving student achievement in business programs: The effective use of standardized assessment tests. *Journal of Education for Business, 79* (2), 90-98.

Braun, N. M. (2004, Mar/Apr). Critical thinking in the business curriculum. *Journal of Education for Business 79* (4), 232-236.

Eldenburg, L. E. & Wolcott, S. K. (2005). Hudziak Industries (Problem 9.32). Cost management: Measuring, monitoring, and motivating performance. Hoboken, NJ: John Wiley & Sons.

Erwin, T.D. (1991). *Assessing student learning and development: A guide to the principles, goals, and methods of determining college outcomes.* San Francisco, CA: Jossey-Bass.

Erwin, T. D. for the U.S. Department of Education, National Center for Education Statistics. (2000). *The NPEC Sourcebook on Assessment, Volume 1: Definitions and Assessment Methods for Critical Thinking, Problem Solving, and Writing.* NCES 2000—172, Council of the National Postsecondary Education Cooperative Student Outcomes Pilot Working Group: Cognitive and Intellectual Development. Washington, DC: U.S. Government Printing Office.

Gainen, J., & P. Locatelli. (1995). *Assessment for the new curriculum: A guide for professional accounting programs. Accounting Education Series* (vol. 11). Sarasota, FL: Accounting Education Change Commission and American Accounting Association.

Hofer, B. J., & Pintrich, P. R. (1997, Spring). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research,* pp. 88-140.

King, P. M., & Kitchener, K. S. (1994). *Developing reflective judgment: Understanding and promoting intellectual growth and critical thinking in adolescents and adults.* San Francisco, CA: Jossey-Bass.

Kurfiss, J. G. (1988). *Critical thinking: Theory, research, practice, and possibilities.* ASHE-ERIC Higher Education Report No. 2, Association for the Study of Higher Education.

Lynch, C. L., & Wolcott, S. K. (2001). *Helping your students develop critical thinking skills* (IDEA Paper #37). Manhattan, KS: The IDEA Center. Retrieved data from http://www.idea.ksu.edu/papers/pdf/Idea_Paper_37.pdf.

Palomba, C. A., & T. W. Banta. (1999). *Assessment Essentials.* San Francisco, CA: Jossey-Bass.

Pascarella, E. T., & Terenzini, P. T. (1991). *How college affects students.* San Francisco, CA: Jossey-Bass.

Pellegrino, J., N. Chudowsky, & Glaser R. (Eds.). (2001). *Knowing what students know: The science and design of educational assessment.* Report of the National Research Council, Center for Education, Division on Behavioral and Social Sciences and Education. Washington, D. C.: National Academy Press.

Walvoord, B. E., & V. J. Anderson. (1998). *Effective grading: A tool for learning and assessment.* San Francisco, CA: Jossey-Bass.

Wolcott, S. K. (2002). *Competency assessment methods.* Assessment guidance in the library of the AICPA Educational Competency Assessment Web site. New York, NY: AICPA. Retrieved data, from http://eca.aicpaservices.org/.

154

Wolcott, S. K., & C. L. Lynch. (2002). *Developing critical thinking skills: The key to professional competencies.* Sarasota, FL: American Accounting Association.

Wolcott, S. K., C. P. Baril, B. M. Cunningham, D. R. Fordham, & K. St. Pierre. (2002). Critical thought on critical thinking research. *Journal of Accounting Education, 20*, 85-103.

Author Bio

Susan Wolcott earned her Ph.D. from Northwestern University in Accounting and Information Systems. She has written extensively in the area of critical thinking and has conducted numerous workshops on the topic. She is currently the owner of WolcottLynch Associates, which specializes in innovative and practical ways to enhance and assess critical thinking and professional problem solving skills. She has held faculty positions at The University of Denver, University of Washington, and Helsinki School of Economics. She is a co-author of *Cost Management: Measuring, Monitoring, and Motivating Performance* (New York: John Wiley & Sons, 2004).

If you were to ask business school deans, the business community, and the public "What is the most important issue related to business curriculum today?" a popular response would be "business ethics." In fact, the editors' survey of business school deans found "ethics" to be one of the top two learning goals for business programs. There is controversy about how to teach ethics and whether it can be taught at all – however, given today's climate, it is a curriculum challenge that must be met. A complicating factor is that "ethics" can mean very different things to different people. Furthermore, it is difficult to assess. In the chapter below, the author provides useful insight regarding how to define and assess this complicated, and increasing relevant, learning goal.

CHAPTER 9
LEARNING IN BUSINESS ETHICS COURSES:
INITIAL IDEAS ABOUT CONTENT AND ASSESSMENT
Timothy J. Fogarty
Case Western Reserve University
Weatherhead School of Management

Ethics as a Learning Goal

In most areas, business education has crystallized an agreement about what should be taught. This degree of consensus, albeit imperfect, facilitates the assessment of learning outcomes. Although assessment is never easy, it becomes much more difficult in an environment of content confusion and process problems. This chapter pertains to one of these difficult content areas.

In the minds of external stakeholders, few subjects rival business ethics as an important element of postsecondary education (Cohen & Pant, 1991). A survey by the editors of this volume also shows that business school deans rank ethics among the top five learning goals for their programs. Recent scandals pertaining to malfeasance in corporate governance exacerbates the need for universities to marshal resources in this area. The need for business people to have a well developed moral sensitivity and to exhibit informed ethical judgment resides at the core of this priority. In a sense, therefore, such moral sensitivity and informed ethical judgment are essential outcomes for business programs.

Notwithstanding the need for curriculum initiatives to offer learning opportunities in business ethics, progress has been slow and sporadic. Two fundamental hurdles constrain movement in this area. First, pursuit of this learning goal has been countered by those who argue that ethics cannot be meaningfully taught in the classroom. Although this position has in some

quarters been made moot by legislation and accreditation guidelines, its persistence in the literature (e.g., Thornton & Peterson, 2000) illustrates the problematic status of its pedagogy and the lack of consensus about its contents. A second hurdle relates to the bundling of ethical materials and ambivalence about where and how ethics students should learn about ethics. Whereas the usual solution of a dedicated, stand-alone course exists, the idea that ethics should be sprinkled across many courses continues as an alternative (Kenny & Eining, 1996).

This paper accepts as a maintained hypothesis that business ethics can and should be taught in college. A position on whether ethics should be a separate element of the curriculum is not taken. Perhaps, business school faculties should consider desired outcomes as an essential part of that decision.

Because ethical education is so intensely contested, this chapter allocates some preliminary consideration to some normative objectives of teaching and learning in the area. This requires explicitness about what business ethics education is not about. Once these goals are established, the central dimensionality of a normative pedagogy can be described. Only at this point can ideas about measurement and assessment of educational outcomes be offered. Thus, an unusually high degree of demystification is essential.

Ethics Education: What it is and is Not

One reason that ethics sits so uncomfortably within the business curriculum is its opposition to the telos of the remainder of the coursework. Whereas the latter is individualistic, acquisitive and technical, the former should be oriented toward the collective good, be purposefully cautious of wealth as the ultimate "bottom line" and not be reduced to a set of calculations or rules. Teaching and learning in the area of ethics, therefore, strikes many as a false note in the larger symphony of capitalism. By putting profit-making in its social context, learning about ethics demands that the student go further than other areas in the business curriculum, such as business law, which define property rights and identify non-owner stakeholders.

Students are made uncomfortable by lessons that devolve to varying ideas of right and wrong, even when the origins of both are well explored. Perhaps reflective of an educational system that too predictably provides correct answers, students prefer bright lines of demarcation rather than the elicitation of situations and issues that can, at best, provide a "feel for the game." Nonetheless, a curriculum that successfully challenges students to appreciate how shared ideas of appropriate behavior are maintained would accomplish a valued objective. If students can go from believing in a natural law of the marketplace to understanding the normative and the political systems needed to keep a market in place, the quality of the experience is difficult to gainsay. A curriculum that engages students in active and passionate

conversation about such issues allows them to identify the artifactual ideas and the thin espousals embedded in other arguments. Students who can acknowledge that reasonable people can still disagree about the correct course to follow are making distinct progress as young professionals.

Before ethics can be approached, educators must disabuse themselves of unproductive ideas that are unfortunately intertwined with ethical education. First, ethics need to be disassociated with virtue. Whereas it may be that people cannot be made virtuous, they can be taught ethics. Virtue overlays excessive baggage, some of which emanates from organized religion, about the struggle between absolute good and absolute evil. As such, virtue is best bracketed and put to the side. Along similar lines, a good ethics class should not be captured by philosophy. Albeit regrettable, the business student is likely to "check out" if the theme becomes "what would Aristotle do?" Teaching and learning about ethics should be very much integrated with any particular day's *Wall Street Journal*. However, alternative misdirections will occur if a corpus of rules becomes the centerpiece. Too often, ethics in the business curriculum is viewed as an opportunity to drill into students' minds chapter and verse of various codes of conduct either directly or indirectly through the provision of a litany of clear lapses. Such an orientation instills a proceduralism that encourages the pursuit of loopholes (Vyakarnam *et al.*, 1996) and de-emphasizes the development of judgment (Pincus, 2000). This approach becomes even more static when students are asked to examine the enforcement record of governments or professions. This emphasis leads to a dysfunctional focus on detection rates (Wallace & Wolfe, 1995) and concerns over sanction equity (Moriarity, 2000). The failure to agree with the eschewing of these paths will make the proposed assessment program below less valuable.

Overall, ethics in the business curriculum must distinguish between teaching about virtue and offering opportunities for students to develop moral sensitivity and the capacity to make informed ethical judgments. Ethics in the business curriculum must also offer students an opportunity to understand how shared ideas of appropriate behavior in business are developed and maintained.

Dimensions of Ethics

To some, every action has ethical overtones. However, few would deny that some situations present more obvious and more consequential ethical content. Therefore, the first threshold dimension for students is moral intensity. This entails detailing the consequences of business action so that the interconnections between parties can be explicit, and the idea of "no harm, no foul" can be diffused.

Perhaps more broadly, educators often underestimate the variation that exists in the moral sensitivity of students. Even within situations rife with potential consequence, some students exhibit little consideration for

the concerns of others. The corporate form, in which a hypothetical entity is the official actor, allows students to become business practitioners unaware of other perspectives on their actions. Ethical sensitivity appears to be a separable dimension worthy of pursuit as an objective of business education (Burns *et al.*, 1999).

Underlying ethical behavior lies a well-developed conception of fairness. Contrary to the predictions of agency theory, self-interested behavior does not necessarily preempt the field. In practice, organizational fairness ranks high in the criteria for the most desired positions (Trevino & Nelson, 1995). Managers may be willing to sacrifice personal rewards in the short run for allocative fairness in the longer term (Pant *et al.*, 2001). Students need help in learning to differentiate equity and equality.

Building upon the previous dimensions, faculty members should seek to provide students with a sense of accountability. Although in a literal sense, this term references adherence to the system of controls in place to prevent untrustworthiness, the spirit of accountability transcends these instrumental devices and suggests a transparency to others that furthers legitimate interests. Unlike other ethical dimensions, accountability lacks strong parallels in non-business applications and requires strong formal and informal socialization to be nurtured.

Business ethics also point to a certain business professionalism that make them unique. Ethicality is a cornerstone of the attitude that business should leverage superior competence toward the public interest. While professionalism may be a broader and more nebulous construct than ethicality, the confluence suggests that it should be treated as a dimension that is sought. Like accountability, professionalism requires explicit consideration of how people react in the presence of strong rewards that are not always perfectly aligned with ethical conduct.

Substantive Areas
The domain of business ethics is sufficiently rich and diverse to preclude attempts to inventory or catalogue the cases that could be used. Virtually no business subject is without its ethical overlay. The classic ones exhibit the overlap with marketing wherein appeals to customers might utilize the baser instincts or where promotional efforts might include incentives designed to tempt others to compromise their integrity. In that every company needs to protect certain transformation information, ethical decisions also could be cast as production issues with a technological core. In the modern economy, informational integrity is essential, thus rendering management information systems into the sandbox for people of all ethical stripes. For example, managers have long been suspected of creating budgetary slack for their units by injecting misleading information into the process (Douglas & Weir, 1999).

In response to the recent crisis of faith in some large corporations,

accounting provides some unusually fertile applications. High ranking managers have the ability to "manage" earnings, a practice that satisfies many corporate constituents. The ethicality of such practices could be vigorously debated (Sennetti *et al.*, 2003). Increasingly, the entirety of corporate disclosure has been addressed in ethical terms. Often, this entails information about the infliction of harm to the physical environment (e.g., Adams, 2002) but could also embrace issues of importance to employees.

Because all organizations have difficulties to overcome, the ethical template is often affixed to the ways in which the issues are handled. The circumstances of the whistleblower continue to provide virtually endless debate about preferred actions (Pamerlee *et al.*, 1982; Brabeck, 1984). More systematically, the relationships between high corporate officials (including boards of directors) and internal and external watchdogs (auditors) need to be understood as ethical contests (Mitschow & Langford, 2000; Doucet & Eprile, 2000). The objective of pinpointing these relationships is to understand that certain reactions such as retaliation and exploitation of a superior informational advantage can compound the ethical difficulty.

Only in the last few years has there been a systematic effort to embrace the topic of fraud as a business issue. The criminality of fraud does not excuse it from the tent of organizational studies. Ethical study cannot be limited to encouraging people to do the right thing. It must also incorporate active awareness that "the dark side" lurks everywhere that intention, opportunity and rationalization cohabitate. Students with higher levels of moral reasoning ability are better able to detect fraud risk areas (Roberts & Koeplin, 2000).

Pedagogy

The nature of the business ethics field tends to preclude the use of many of the practices that are the stock in trade of business education. Only with great difficulty could one imagine that a lecture-based course could be much of a success. In addition to the critical lack of "facts," the ethics area is challenged to describe much of value in an organized, didactic way.[48] Ethics seems to demand the hands-on engagement of the mind that is more likely to result from highly interactive teaching choices. Students in this mode should be asked to convert ethical ideas into sound action choices, rather than to hear accounts of those that have and have not.

The case method would seem to be ideally suited to the ethics course. Cases would seem to offer a robust set of facts that would best be able to place ethical choice in a suitably complex swirl of contingencies. Cases also provide students with sufficient factual backgrounds to convince the reader that the events could be real, *Ceteris parabus*, realism makes it more likely that students will invest themselves in resolution. With the case model, students are also invited to step outside the disciplinary silos that their major fields of study already have begun to construct for their thinking about issues.

More than ample materials exist to support the case method in ethics. A large number of cases, most drawn from the business world of the last ten years, is offered in compilations such as Petrick & Quinn (1997) and Ferrell & Fraedrich (1997). Several groups such as the Hasting Institute offers collections of video case studies.[49] Clearly, the case method instructor of the present day does not have to conjure materials from thin air.

The conversation induced by the case method need not be limited to that which can be contained within the classroom. Organizing students into groups and asking them to take reasoned positions will ensure that the work will continue after class is over. The complexity of the case study often merits the combined effort of students and mimics the fact that many "real world" ethics decisions involve more than a single mind.

The choice of pedagogy is critical in the determination of the range of materials produced by students. A certain affinity exists between the lecture and the multiple choice examination. The case method yields the opportunity to work in a sophisticated way with a large set of materials. Therefore, it provides an excellent opportunity to write essays. Essays, with their open-ended nature, expand the prospects for critical thinking.

The qualitative nature of the material in an ethics course, when matched in character by student work, requires a skillful and dedicated evaluation. Only if this level of work by instructors is forthcoming can the promise of student writing about cases be realized.

Program Assessment

Assessment, stripped of its other connotations, should be understood as the process of systematic evidence gathering motivated by the desire to more fully understand the impact of teaching. It is predicated on the assumption that, since the impact that educational methods has on students is never sufficiently effective, assessment should be used primarily for educational improvement.

This chapter accepts that quality assessment programs begin with the translation of learning goals into an assessment task that can be scored and interpreted. This effort will result in appropriate action, including providing feedback to students and refining the assessment program for subsequent applications. Since best practices in assessment suggest that it be integrated with the actual conduct of the course and that participation be required of students, conformity to those ideals is strongly recommended for ethics subject matter. Other chapters in this book provide more conceptual detail about the general nature of these assessment steps.

Ethics presents a surmountable assessment challenge. Unlike many other business subjects, standard competency batteries have not been produced. The lack of "facticity," already mentioned as a factor in selecting a pedagogy, also weighs in against the pre-test, post-test model of assessing learning. More than just about any other area, the instructor may be staring

161

at the proverbial blank sheet of paper when it comes to designing assessment programs for ethics.

Learning Goals

The use of learning goals as the starting point of ethics assessment implements the valuable advice of beginning with the end in mind. In that learning goals must be few in number, it forces the instructor to explicitly prioritize important outcomes and to let other, less desirable, ones fall more to happenstance.[50]

Although ethics is in many ways *sui generis*,[51] it is also yet another application of the need for students to display the foundational skills of the business profession. Nothing counts unless it can be intelligently communicated. Nothing worthwhile can be communicated in the absence of critical thinking processes. Put into the context of a group, student interpersonal skills come to the forefront. To this trio (knowledge and skills in the domain, communication skills, and critical thinking), one could add the need for students to possess sufficient confidence that their ideas and opinions are valuable. Learning goals in ethics should be designed so that these skills can be clearly observed.

Earlier in this chapter, several dimensions of ethics were extracted as a means to define the scope of the subject. These ideas can also help shape the conceptual goals of assessment. At the highest level of abstraction, ethical awareness can be paired with ethical reasoning as the twin peaks of the ethics class. Some might be interested in pairing moral sensitivity and moral intensity as legitimate categories of ethical awareness. Along similar lines, ethical reasoning may prove an excessively aggregated construct. Accordingly, it could be reasonably partitioned into fairness definitions, professionalism and accountability. The latter two subsets are especially consequential to the business student, and represent content areas usually not found in philosophic renditions of the ethics subject.

At a more operational level, an instructor may desire to consider the following objectives:

1. Getting the facts straight.
2. Extrapolating the consequence of action to other parties (including those remote from the action).
3. Assessing utilitarian trade-offs among the parties involved.
4. Understanding the hierarchy of rights recognized by society.
5. Appreciating how varying conceptions of equity can result in different evaluative conclusions.
6. Looking at transactions from the perspective of other parties, including corporate actors.
7. Adjusting for generic biases known to exist in decision making.
8. Learning from history.

9. Limiting the expression of self-interest and other outcomes of marketplace logic.
10. Abstaining from the tendency to justify the means by virtue of the end.

These outcomes provide a larger set of assessment traits within the broader categories extracted in the previous paragraph. In turn, they also could be divided and rearticulated. However, no list is definitive. Every list of objectives should be considered a work in process and the primary area for subsequent improvement.

General Measurement Methods

For many, any discussion of ethics requires the Defining Issues Test (DIT). This measure has dominated the U.S. literature and acts as the lens through which many understand the field. It purports that individuals populate a series of stages that indicate their level of ethical development. Accordingly, the DIT facilitates a snapshot comparison of various groups in a very simple way. Although the DIT has no shortage of validity and reliability issues (Bay, 2002), its use has sufficient recognition so as to provide instant credibility. Ironically, the DIT has not been used enough as an assessment device. Properly interpreted, a pre- and post-course administration might be insightful.

For many, the DIT draws upon a shopworn, albeit classical conception of ethics. Many would prefer a scale that draws more upon social responsibility. This may be accommodated with the Multidimensional Ethics Scale (MES). This instrument more strongly features moral equity, relativism and contractionism dimensions of the construct. Some evidence also suggests a superior linkage to behavior (Henderson & Kaplan, 2002).

The important idea for educators to take away from the general measurement issues is that although perfect assessment methods do not exist, reasonably good ones have considerable value. In addition to the range of test phenomena, educators should consider the difference between the prescriptive and the descriptive levels. Whereas we might wish that there would be a merger between what people should do and what people will do, empirical evidence runs toward the contrary conclusion (e.g., Thorne, 2000). In business applications, this measurement problem is increased by the varying degree that students appreciate and embrace an appropriate value orientation. Last, unlike other subjects, ethical intentions are confounded by social desirability bias. Whereas not everyone agrees with the correct course of conduct, some will profess a more desirable path not because they believe it, but they believe that is what those administering measures wish to hear. Various corrective procedures have been suggested (e.g., Geiger & O'Connell, 2000).

The limitations of established measurement suggest that assessment cannot be based in any material part of "off the shelf" solutions. Instructors, however reluctantly, must take out a fresh piece of paper. The task is not to

take a one-time reading of the level of student overall ethicality, but to measure their appreciation for the process of becoming an ethical actor. Established measures may provide students something to be critical about more than they provide insight.

Goal-Oriented Measurement Methods

Although measurement sounds like a highly mechanical process, it imbeds many important instructor choices. The discussion of pedagogy above argued that the case method was uniquely applicable to the ethics subject matter. Accordingly, ideas for evaluation are limited to those within this overarching method for soliciting student deliberation and choice. Furthermore, any ideas for measurement must be mindful of the learning goals and specific objectives of a program.

Imagine a series of hypothetical cases, each designed to focus upon a narrow set of learning objectives as outlined in this chapter. Students, either individually or grouped into teams, would confront highly focused questions that would squarely frame the objective. Students would be asked for their logic and commanded to persuade a skeptical audience that the position they took was better than others. The point of this would be to observe the extent to which a student could synthesize the elements of ethicality and produce a decision that optimally brought these to bear.

The subject matter of the cases would have to possess the look and feel of authenticity, yet not be real. Cases could be based on the substantive ethical issues discussed earlier in the chapter. This could be achieved by using a historical case as a skeleton around which a set of more suggestive and tantalizing facts could be hung. This would enable the analytical separation of a particular dimension in a more distinct manner than the real world allows.

The standards of evaluation should be rigorous and absolute. Students should not be credited with that which they meant or that which could be inferred from their words. The ability to establish correct causal sequences and to appreciate the likely behavior of others is an important competency to score. Instructors should avoid the temptation to become excessively relative with regards to the range of acceptable response. The difficulty pertains to the position of the line that separates acceptable and non-acceptable student response. Here, some degree of subjectivity cannot be avoided. A predefined rubric that includes a set of traits and two to three performance levels (acceptable/unacceptable; good/fair/poor) would facilitate the scoring process. Other chapters in this volume more fully discuss the value and use of assessment rubrics. Table 1 presents an example for the ethics area.

Instructors should find value in implementing multiple measures in the scoring process. In addition to possessing a finely-toothed rubric, instructors may want to deploy a second reader. Instructors never are able to see everything, no matter how good they are in evaluation mode. If evaluation can be decoupled from grading, instructors can find creative ways for students

Table 1
An Example of a Rubric to Assess Ethics*

Traits	Good	Fair	Poor
1. Correctly identifies the facts and ethical issues.	Shows good appreciation for the facts and the ethical issues involved	Shows reasonable appreciation for the facts and ethical issues	Fails to show an appreciation for the facts and ethical issues
2. Extrapolates the consequence of action to other parties.	Recognizes all stakeholders and fully explores consequences	Recognizes stakeholders and explores some of the consequences fairly effectively	Fails to recognize all stakeholders or explores consequences ineffectively
3. Assesses utilitarian trade-offs among the parties involved.	Assesses key trade-offs and does so effectively	Assesses some of the trade-offs and does so fairly effectively	Fails to assess trade-offs or does so ineffectively
4. Understands the hierarchy of rights	Recognizes and appreciates the hierarchy of rights	Recognizes and shows reasonable appreciation for some of the hierarchy of rights	Shows no understanding of the hierarchy of rights
5. Appreciates how varying conceptions of equity can result in different evaluative conclusions.	Effectively demonstrates appreciation for different conceptions of equity	Demonstrates fair appreciation for different conceptions of equity	Shows no appreciation for different conceptions of equity
6. Looks at transactions from the perspective of other parties, including corporate ones.	Effectively recognizes the perspectives of other parties	Fairly adequately recognizes the perspectives of others	Fails to recognize the perspectives of others or does so poorly
7. Recognizes own personal biases that can influence decision making outcomes.	States assumptions and identifies and clarifies personal beliefs that may affect decision outcomes	States assumptions and identifies but does not clarify personal beliefs that may affect decision outcomes	Does not state assumptions or does not identify personal beliefs that may affect decision outcomes
8. Learns from history.	Demonstrates a good appreciation for prior history where managers faced similar ethical issues	Demonstrates a fair appreciation for prior history where managers faced similar ethical issues	Does not demonstrate an appreciation for prior history where managers faced similar ethical issues
9. Limits the expression of self-interest and other outcomes of marketplace logic.	Demonstrates substantive constraint in the expression of self interest and other outcomes of market-place logic	Demonstrates some constraint in the expression of self interest and other outcomes of market-place logic	Demonstrates no constraint in the expression of self interest and other outcomes of market-place logic
10. Abstains from the tendency to justify the means by virtue of the end.	Clearly abstains from tendency to justify the means by virtue of the end. Issues are clearly considered and decision-making is mindful of this tendency	Shows awareness of the tendency to justify the means by virtue of the end, but actions to avoid the tendency are not very deliberate	Does not abstain from the tendency to justify the means by virtue of the end

* This rubric was designed for illustration purposes only. Normally a rubric that will be used for program assessment incorporates the expectations for students' learning based on the program's mission and the shared values of faculty who teach in a program.

to participate in the evaluation. Student reflective self-criticism should also not be ruled out.

Interpretation is the true instructive art. Whereas the comparison between a student's answer and the instructor's idea of a perfect answer can be cognitively demanding, it pales in comparison to appreciating the nature of the student's flawed reasoning. Here the instructor has to make suppositions that might be more speculative than conclusive. These should not be over generous, however, since the burden of proof has to lie upon the student. Often interpretation requires that the case assessment be read holistically, so as not to confuse poor organization with lack of knowledge.

Ethics instructors will encounter the occasional response that is both intelligent and creative. "Outside the box" thinking should be highly praised if new ways to find predictable dilemmas can be isolated. Some tolerance for debating the presumptions of the question should be extended if it leads to valuable insights.

165

Implementation and Closing the Loop

Assessment is only as good as it is practiced. To be practiced, assessment must have influence. To have influence, it must not be static. All of these actions fall within the infamous "closing the loop" step of assessment, a step that has proven much more difficult than the others.

Implementation digs into the more detailed ways instructors plan and manage assessment. Since assessment cannot be divorced from the ongoing flow of the curriculum, specific details cannot be generalized. Nonetheless, some general ideas can be quite powerful in the shaping of the processes that make up implementation.

The best wisdom about implementation is contained in the classic expression, "If you don't want to know, don't ask." This suggests that information that is collected has to be used. Evaluation of a student case analysis should be sufficiently focused so that a conclusion can be drawn about a specific competency. Thus, assessment exercises should not be fishing expeditions that produce information that cannot be used or is irrelevant to the purpose.

A good deal of assessment information must be used in the feedback process. Letting students know how they have done, with the hope that they can do better in the future, is a greatly underappreciated element of assessment. Without constructive feedback, assessment threatens to be external to the lives of students and only of benefit to educational institutions. Feedback is the only device that has the potential to lead to changed behavior by the students at hand. Whereas the value of inter-cohort course changes should not be ignored, the best practices are those that treat every set of students as the work in progress. The value of feedback is strongly conditioned by its appropriate timing.

While ethics cases with particular learning objectives should be spread across the semester, there should also be a capstone type event that integrates the various dimensions of ethics and fully explores achievement of the learning goals for ethics in the curriculum. This can be implemented through the familiar case competition vehicle. Schools might consider using external judges (perhaps advisory board members) for such an event.

More important than the timing and judging of case events in the ethics class is the learning that the instructor can accomplish from one offering to the next. Evidence of imperfect communication between the facts of the cases and the responses of the students should trigger alterations of the cases, such that the ethical points can be better elicited. Such changes will prevent cases from becoming stale to the instructor, or compromised by virtue of the "grapevine" among students. The instructor needs to learn the lessons of history with the materials.

Pitfalls

No faculty member has ever been short of several reasons not to do

166

something new. Since assessment beyond the surface of the grades given is relatively new, pitfalls abound and need to be anticipated.

One set of problems relates to trying to do too much. We should never forget that the perfect is the enemy of the good. Just as there can be no perfect assessment program, there can be many different good ones. One of the ways that excessive perfectionism creeps into assessment is through an excessive concern with validity. We should not be importing the standards of our research into our assessment efforts unless we want most of our work rejected.

At the same time, we should not make assessment too narrow by failing to be open to supplemental or alternative ideas. The "not invented here" syndrome prevents crediting others with good ideas than can be successfully imported. There are more similarities than there are differences in the interfaces between students and faculty no matter how special we think we are.

There will always be some pressure to pull assessment back to the simple measures that we have accepted for many years. These need to be avoided. For example, student evaluation of teaching, which tends to express not much more than satisfaction, is an excessively crude instrument that may be exceptionally ill-suited to the ethics area. Likewise, exit interviews and the polling of employers exhibits a *post hoc* response inconsistent with the spirit of assessment.

The last pitfall that should be mentioned is a tendency toward zero-sum thinking. Assessment that is properly designed should benefit all who are affected by it. Armed with the knowledge that there is something of value for all, participants should not be paralyzed by temporary imbalances in *cui bono*.

The Larger Picture

For many organizations of higher learning, ethics are central to mission statements. We need to not just look at these expressions as so much fluff. Proof that a school is successful at teaching ethics would be a very convincing part of the organization's overall value to society.

The successful assessment of ethics could be a centerpiece in a larger attempt to come to grips with the quality of instructional effort. If a course like ethics can be assessed, any course can be assessed. Seeing this proof may make the gospel of continuous improvement more palatable to those that would otherwise resist.

Innovations in assessment provide excellent opportunities to contribute to the scholarship of teaching and learning. Many natural experiments become possible when the data produced by these efforts are harvested and put into the framework of hypotheses.

Assessment "Noise"

Any assessment system is only as good as the user's recognition of

167

extraneous influence. Just as there are always unspecified variables that contribute to the constant in a regression equation, a miscellaneous set of other influences has to be considered in educational assessment.

Countless ways have been devised by those trained in psychology to describe personality differences. Personality refers to an organized means of response to situations. Any instructor who has vigorously interacted with students would attest to the influence of personality. If personality is a meaningful construct, it should offer some resistance to change. This presents some difficulties because, in an ethics class, we may be expecting a systemic change. For example, Uddin and Gillett (2002) find that multiple dimensions of personality are significant predictors of ethical disposition. This problem may not be as extreme in other areas of the curriculum, because the mastery of a set of facts in the acquisition of a skill may not have to engage student personality as directly.

Men and women have long been suspected to be more different than similar as ethical actors. Almost without exception, women have been shown to have more acute and well developed ethical abilities, however measured (e.g., Etherington & Schulting, 1995). Even so, gender differences may be context specific. For example, males seem more willing to penalize ethical infractions (Singa-Mugan & Onkal-Atay, 2000). Nonetheless, gender is a force sufficiently powerful to demand attention in any assessment effort.

Recent demographic shifts have increased our appreciation for national origin and race as other sources of influence. The line that separates acceptable and unacceptable business behavior is drawn by cultures in different places. Thus, teaching and learning about ethics and, in particular implementing what is learned, can be challenging (Siegal *et al.*, 1997). Often, Hofstede's work provides some broad expectations for the major vectors of variation between people from different countries (e.g., Brody *et al.*, 1999). Unaccounted for, these influences may disguise the actual appreciation for ethical thinking.

Conclusion

When business faculty are asked about the ideal graduates they would like to produce, they invariably gravitate toward ethicality as a central attribute. This desire cannot be left to chance to occur. Instead, it illustrates the need to assess the success of our efforts in this educational area.

Some degree of humility is appropriate as we walk toward this task. We would like to produce more ethical managers, and as a result influence the social responsibility of corporations. Our ability to reach behavior with education is small in this area (Lane *et al.*, 1988). The ethical ideas and behaviors learned in the best programs can be seriously compromised by the passage of time and the contrary socialization offered by some employers. Nonetheless, our task is to do better than we have before.

Teaching and learning in the ethics area is subject to considerable variance in its main pedagogical direction and its primary objectives. Naturally,

168

these will have consequence for its assessment. A balance between extremes is called for in curriculum design and in the selection of materials. Educators would do well to remember that positive information is more likely to be used for decision making than is negative information (Hoveland & Weiss, 1953). That which our students learn must be more than knowledge about the infamous ethical shortfalls of business history.

This chapter has attempted to approach ethical instruction and assessment as a conversation. Here it is assumed that students learn by talking and by writing. Within the conflict created by simulated action, they become aware of choices and their consequences. Thus, the practice of ethics is a skill that must be acquired. Only then ethics can have emancipatory power.

Assessment is one of several vanguards that are reconstructing the nature of the business academy. In this position, much time will need to elapse before assessment is a natural part of higher education. Until then, many faculty will do it reluctantly or under coercion (Fogarty, 2004). However, assessment is here to stay and will eventually be recognized as the right thing to do. This developmental, or perhaps embryonic, state also aptly describes ethical education. Ethical education is much more than teaching and learning about business ethics. It involves moral sensitivity and ethical judgments about the structure and content of our curricula and programs. On our way toward the realization that all education is moral, we ought to have highly effective ethics instruction that puts these issues in the crosshairs.

Endnotes

48 This conclusion is predicated on the acceptance of the definition of ethics discussed earlier in the chapter. Instructors who wish to, for example, teach codes of ethics could find lecture techniques more suitable.

49 Given the increasing difficulty of students to extract information from written texts, offering visual cases may be highly effective.

50 Outcomes that are made visible and made the subject of measurement will be more likely to be produced by students than those that are inchoate.

51 In addition to not fitting within a disciplinary field, ethics is not usually associated with the sort of skill development that can be readily packaged and sold in employment markets.

References

Adams, C. (2002). *A Critique of Reporting on Ethical Issues: The Case of ICI.* Unpublished paper, University of Glasgow.

Bay, D. (2002). A critical evaluation of the use of the DIT in accounting ethics research" *Critical Perspectives on Accounting Ethics,* 13, 159-177.

Brabeck, M. 1984. Ethical characteristics of whistle-blowers. *Journal of Research in Personality, 18,* 41-53.

Brody, R., Coulter, J. & Lin, S. (1999). The effect of national culture on ethical decision making. *Teaching Business Ethics, 3*(4), 383-398.

Burns, J., Kiecker, P. & Spindle, R. (1999). *Ethical sensitivity: Its conceptualization and investigation among tax practitioners.* Unpublished paper, Texas A & M University.

Cohen, J. & Pant, L. (1991). Beyond bean counting: Establishing high ethical standards in the public accounting profession. *Journal of Business Ethics, 10*, 45-56.

Doucet, M. & Eprile, B. (2000). The public interest: Ethical issues facing external auditors and financial reporting regulators. *Research on Accounting Ethics, 6*, 221-242.

Douglas, P. & Weir, B. (1999). *Integrating ethical dimension into a model of budgetary slack creation.* Paper presented at the annual meetings of the American Accounting Association, San Diego.

Etherington, L. & Schulting, L. (1995). Ethical development of accountants: The case of Canadian certified management accountants *Research on Accounting Ethics, 1*, 235-251.

Ferrell, O. & Fraedrich, J. (1997). *Business Ethics: Ethical decision making and cases.* Boston: Houghton-Mifflin.

Fogarty, T. (2004). Why faculty resist assessment. In T. Calderon and B. Green (Eds.), *Best practices in accounting program assessment.* Orlando: Teaching & Curriculum Section, American Accounting Association.

Geiger, M. & O'Connell, B. (2000). An examination of using surrogate measures to access social desirability bias. *Research in Accounting Ethics, 6,* 107-127.

Henderson, B. & Kaplan, S. (2002). *An Examination of the Role of Ethics in Tax Compliance Decisions.* Paper presented at the Ethics and Professionalism Conference, San Antonio.

Hoveland, C. & Weiss, W. (1953). Transmission of information concerning concepts through positive and negative instances. *Journal of Experimental Psychology, 45*(3), 175-182.

Kenny, S. & Eining, M. (1996). Integrating ethics into intermediate accounting. *Research on Accounting Ethics, 2*, 159-186.

Lane, M., Schaupp, D. & Parsons, B. (1988). Pygmalian effect: An issue for business education and ethics. *Journal of Business Ethics, 7*, 223-229.

Mitschow, M. & Langford, D. (2000). Whistleblowing by independent auditors. *Research on Accounting Ethics, 6*, 91-106.

Moriarity, S. (2000). Trends in ethical sanctions within the accounting profession. *Accounting Horizons, 14*(4), 427-439.

Pamerlee, M., Near, J. & Jensen, T. (1982). Correlates of whistle-blowers' perceptions of organizational retaliation. *Administrative Science Quarterly, 27,* 17-34.

Pant, L., Sharp, D. & Cohen, J. (2001). *The effect of perceived fairness and moral development in an agency context.* Paper presented at the annual meetings of the American Accounting Association, Atlanta.

Petrick, J. & Quinn, J. (1997). *Management ethics: Integrity at work.* Thousand Islands, CA: Sage.

Pincus, K. (2000). The role of rules in accounting: Do rules guide us to the right path or to the path of least resistance? *Research on Accounting Ethics, 6*, 243-258.

Roberts, D. and Koeplin, J. (2000). *The impact of ability to perform moral reasoning on knowledge acquisition and fraud risk assessment.* Paper presented at the Ethics and Professionalism Conference, American Accounting Association, Philadelphia.

Sennetti, J., Shawver, T. & Bancroft, P. (2003). The moral and cultural reasoning of IPO accountants. *Research in Accounting Ethics, 9*,141-167.

Siegal, P., Mosca, J. & Karim, K. (1997). The role of cultural differences in implementing a code of ethics. *Journal of Accounting Ethics and Public Policy.*

Simga-Mugan, C. & Ontzal-Atay, D. (2000). *Are there gender differences in ethical judgment: Accounting versus general business settings.* Paper presented at the Ethics and Independence meeting, Philadelphia.

Thorne, L. (2000). The development of two measures to assess accountants' prescriptive and deliberative moral reasoning. *Behavioral Research in Accounting 12*, 139-170.

Thornton, J. & Peterson, B. (2000). Codes of conduct and ethical perceptions: A Comparison of accounting and military students. Paper presented at the Ethics and Independence Conference, Philadelphia.

Trevino, L. & Nelson, K. 1995. *Managing business ethics: Straight talk about how to do it right.* New York: John Wiley.

Uddin, N. & Gillett, P. 2002. The effects of moral reasoning and self-monitoring on CFO intentions to report fraudulently on financial statements. *Journal of Business Ethics, 40*(4),405-407.

Vyakarnam, S., Srikanthan, S. & Fitzsimons, S. (1996). Can accountants distinguish their assets from their morals? *Business Ethics: A European Review, 5*(3),156-163.

Wallace, W. and Wolfe, C. (1995). Do expected audit procedures prompt more ethical behavior? *Research on Accounting Ethics, 1*,145-167.

Author Bio

Timothy J. Fogarty is the KPMG Peat Marwick Faculty Fellow professor in the Accountancy Department at the Weatherhead School of Management, Case Western Reserve University. He has published over one hundred twenty articles on a wide variety of topics, in both academic and practitioner journals. His research interests include accounting education, the sociology of business organizations, and the regulation of professionals. He has recently been involved with the accreditation of business programs at universities for the AACSB and the CPA examination for the AICPA. At Case Western, he currently serves as department chair, teaches auditing, income tax and the legal environment of management, and serves in several university-level capacities.

We have included this topic, in part, as an illustration of how an ineffable learning goal can be assessed routinely in a classroom environment. Drawing from the literature to define creativity, the author translates it into observable traits that can be evaluated through classroom exercises and assignments. Furthermore, the detailed discussion of "closing the loop" provides suggestions on how creativity can be promoted in business school curriculum. With entrepreneurship fueling much of the growth in the US economy, and creativity cited as a critical success factor in a growing number of industries, creative thinking will likely emerge as a learning goal for many business schools in the future.

CHAPTER 10
CREATIVITY: THE ALL PURPOSE BUSINESS TOOL (AND YES, IT'S ASSESSABLE TOO!)

Joseph Aniello
Francis Marion University

Introduction

As small businesses become more and more important to America's role in the 21st century global economy, business schools are faced with the new and different challenges of better preparing students for their future. With so many technological advances during the prior generation of business evolution, the areas of communication, transportation and information have become cheaper, more reliable, faster and more powerful than ever before. These factors have made high-speed, far-reaching world market participation both accessible and affordable to entrepreneurs anytime and anywhere.

Just as technology has provided many of the new and bigger opportunities for business during the last wave of new venture development, there is an additional essential ingredient necessary for the continued growth of the business discipline. Creativity is emerging as the new technology that will be critical to continued success for business people and in turn, crucial for the competitiveness of business schools. At the risk of paraphrasing the Old Testament, creativity begets innovation, and innovation begets entrepreneurship. If entrepreneurship will be the vehicle of the new economy, then innovation will be its engine, and creativity is the fuel that powers that engine. The most pertinent question then becomes, will that fuel be "regular" or "high-octane? "

Unfortunately, many people still maintain misconceptions about creativity as it pertains to the business discipline and its related activities. Too often, the definition is confused with the notion of talent as it is associated

with artistic endeavors. While talented people are almost always creative, that creativity is demonstrated through expressions of singing, writing, acting, dancing, painting, etc. In a much broader sense, however, creativity can be described as "generating idea alternatives for problem solving." As any small business owner will undoubtedly attest, that need is a daily occurrence in the performance of his or her working roles and responsibilities. A homemaker with small children probably would also agree.

Once comfortable with the idea that not only is it acceptable to combine information in a novel or even unique manner, it follows that this kind of creativity is actually very helpful in achieving ongoing success in the face of business challenges. It can even make a career choice in business more fun, a concept that is often considered antithetical to the mission of business itself. (That would be an entirely different chapter altogether.) Since creativity can be an end unto itself, the challenge for goal-oriented business people is, how does all of this creativity stuff make a positive financial contribution to my bottom line?

When creativity becomes manifested in a tangible product or service, we call it innovation. Just as different fuels are used for drag racers, diesel tractor trailers and passenger cars, creativity is applied differently in order to be considered innovative. In a general sense, all three of the above-mentioned vehicles have engines and use petroleum-derived fuels, but none are interchangeable. In the same way, generally similar processes are used for most creativity, but their specific use is measured by very different success standards depending upon their particular innovation utility.

In the marketplace of business, that utility is most often measured in the competitive terms of profit and loss. Did the customer buy the product or not? Does the consumer like the service or not? Can we sell it for more than it costs us? Is there a demand for repeat purchases? Can we supply that demand in a timely manner? Have we established sufficient barriers to entry? Will our competition threaten our pricing structure? Can we maintain our exclusive distribution channel? Is the advertising compelling enough to prompt action? The answers to these questions and a hundred more reside in the domain of innovation. The creativity of a song may evoke a strong emotional response in a listener, but in business it is an innovation that provokes a strong purchase reaction from the target audience.

That brings us to the "vehicle" of entrepreneurship. If innate creativity has been utilized effectively via the innovation of a product or service, then what happens next? The main difference between inventors and entrepreneurs is that inventors turn an abstract idea into a tangible manifestation of that idea. Entrepreneurs, on the other hand, exploit that tangible manifestation into an opportunity—one which maximizes financial gain. It seems to be at this stage that many business schools begin to aid in the development of the skills necessary to help entrepreneurs grow and manage their emerging ventures.

Many schools now offer courses in entrepreneurship and new venture creation. A "Best Practices" study funded by the Kauffman Foundation includes the following schools as having the best programs for entrepreneurship: Babson College, Baylor University, Carnegie Mellon University, University of Texas, Rensselaer Polytechnic Institute (RPI); University of Colorado, University of Illinois at Chicago, University of Maryland at College Park, and the Wharton School of Business (University of Pennsylvania). The question for consideration is: Do all entrepreneurship curricula offer sufficient emphasis on the creativity/ innovation components of the equation? Without that emphasis, teaching traditional disciplines applied to small businesses is analogous to trying to make a silk purse from a sow's ear.

Entrepreneurship requires many of the same subject matters that exist in larger, more bureaucratic organizations. They must be tailored, however, to the specialized demands and limitations of the small business. For example, with a large marketing budget, most corporate managers can direct those funds toward tried-and-true media vehicles in order to meet the target goals of the overall marketing mix. An entrepreneur, however, very rarely has the resources to run a traditional promotion/advertising/public relations campaign and must rely on tactics more closely associated with a practice called "guerilla marketing." Guerilla marketing relies almost exclusively on creativity; that is to say, using innovative ideas in lieu of money to accomplish the overall objectives. Those objectives can often be the same as the major-expenditure companies; namely, creating more awareness for products and services. A different route must be taken to get there, however.

Similar examples exist for the other subjects taught at most business schools as well. Accounting, economics, finance, management, MIS and anything else that is useful to running a business organization applies to entrepreneurs, just in smaller, less-costly doses. In place of the larger expenditures that big firms can pay for human capital, consultative talent, agency specialization, the latest equipment or the fastest intelligence information, entrepreneurs must rely upon ingenuity to slip under the radar and get there first. Fortunately, due to the small size of most of these firms, they can be more maneuverable, thereby avoiding many of the rules and conventions of the "big wigs." Necessity may be the mother of motorized vehicle invention but without the fuel of creativity, even the most powerful engine would not be able to move the largest or fastest vehicle even one inch.

| CREATIVITY | $\longrightarrow$ | INNOVATION | $\longrightarrow$ | ENTREPRENEURSHIP |

Since the focus of this chapter is on how to increase student creativity and only tangentially innovation and entrepreneurship, we will address some aspects that are foundational. First, here are some definitions that I find

interesting. As it relates to our abilities as people, creativity is "the most complex of all human behaviors" (Runco & Sakamoto, 1999, p. 62). Creativity is to "put people in touch with their values; put people in touch with their purpose and to celebrate diversity" (Lynch & Kordis, 1988, p.137). In even more metaphysical terms, "Life is about invention, not survival. We are here to create, not to defend" (Wheatley & Kellner-Rogers, 1996, p. 11). Paul Torrance, one of the pioneers of creativity study, believed that creativity might be the most distinguishing feature of human behavior which most significantly contributes to excellence in all aspects of life (Torrance, 1974). Thomas Stewart's work concerned itself with how creativity could be utilized in the world of work. He insists that, creativity is the primary raw material and highest order thinking skill of the intellectual capital that will determine an individuals' future worth to society (Stewart, 1997).

The above definitions may be thought-provoking, but in order to assess creativity properly, we will certainly need a more operational definition. It is generally agreed that there are four components of the creative discipline: Person, Process, Product and Press (Isakson, 1988). The rest of this paper will deal with objectives surrounding the assessment of the 4-Ps of creativity and suggested methods for increasing creativity in our students within the context of an environment of higher education. With a strong creative foundation, innovation can take place and then, provide viable opportunities for new ventures.

Assessing Creativity: Objectives, Traits, Standards and Measures

Since creativity will be such a vital skill for success in the business world, one important goal of any relevant business school program should be to help increase a student's creative potential. Quite simply, can students demonstrate more creativity when they graduate from a program than when they began it? This might translate into a learning objective like the following:

Upon completing our program, students will demonstrate more creativity than they did when entering the program as evidenced by their ability to generate multiple, diverse and novel ideas, and to articulate an original idea so others can visualize it.

In essence, this objective calls for classic pre-test/post-test measurement of students' creativity, with their matriculation as the intervention. This may seem like an overly ambitious, non-assessable objective but, as the discussion below will show, it is neither. By the time people reach adulthood, so many biological, psychological and sociological factors have fixed creative abilities at specific range-levels that there is a relatively small amount that remains *un*actualized (Baron, 1969). Accessing each student's own individual latent creativity target, however, might help him/her make a substantial difference in a future competitive situation. The distinction may

be small but then again, usually so is the difference between first place and fourth place in most Olympic events.

As far as measurements against standards are concerned, we can use the widely-accepted definition of creativity's 4-Ps to determine the effect that a business program is having upon the increase of student creativity. The first three—Person, Process, and Product—refer to the student's creative ability and will form the basis of the assessment. The final "P"—Press—refers to the creative environment, and will be discussed in the final section of the chapter when we turn to "closing the loop" remedies.

PERSON

As previously mentioned, there will be a wide range of creativity levels with which students enter any school program. We can not be overly concerned with the absolute measures here (arts programs may), but rather focus on any relative growth that specific students can achieve from their potential to their actual creativity as a result of their academic experiences. In other words, can we release any "latent" creativity prompted by the influence of the teaching environment? A modest goal might be to maximize every student's demonstrable creativity by increasing it 10% over the course of their studies. Naturally, this would require measurements at both the beginning and the end of the program.

There are four dimensions to the measurement of a person's creativity: Fluency; Flexibility, Originality, and Elaboration. These dimensions can form the basis (as traits) to assess students' realized creativity with a simple, timed brainstorming exercise. The method described below will generate data to evaluate all four dimensions and can be easily implemented in a short period of time. (Depending on the presentation format chosen, students could complete this exercise in less than 15 minutes). It could be conducted in or out of class and could be adapted to an online format.

Method: Assign students to think of alternate uses for a common household item like a clothespin, paper clip, wire hanger, drinking straw or masonry brick. In a prescribed amount of time, students list as many ideas as they can, and then prepare a short description of their best idea (either oral or written). Both the list of ideas, and the description of the best idea, is evaluated. If an oral presentation is chosen, students in the class could then respond to a short "quiz" on the presenter's idea to test the audience's comprehension. (A more sophisticated method would be to have the presentations take place in front a panel of local experts, preferably those with experience in new ventures). Written presentations could be evaluated by a panel, the instructor, or a designated rater.

The evaluation criteria (traits) are:

1. Fluency: Defined as the sheer number of ideas that can be generated.

176

Much like fluency in a language—how many words do you know? When it comes to creativity, the more ideas the better.

2. <u>Flexibility</u>: Defined as the variety or diversity of those generated ideas. For example, using a clothespin to seal a snack bag and a trash bag may be two separate ideas but are variations of a theme (same use). Using a clothespin as a picking-up "tool" represents more flexibility of thinking.

3. <u>Originality</u>: Defined as how novel or even unique the ideas are. For example, a macadamia-nut candy bar is only a slightly-original deviation from a peanut candy bar. A raisin candy bar would be a more original modification of an exiting concept. A grasshopper-filled candy bar would be a much more original product idea—even if potentially unappealing. We're not judging yet!

4. <u>Elaboration</u>: Defined as the ability to articulate those ideas. For example, can the idea be described using terms and concepts that people can identify and relate? Or, is the idea too abstract to comprehend?

Possible performance criteria for each of these dimensions (traits) are provided in the rubric below.

After using this rubric to evaluate the list of ideas, the idea the student eventually selects, and the articulation/ presentation of the selected idea, a score can be developed for students reflecting their (realized) creative abilities. Since our learning objective refers to the change in students' creativity, a value added assessment approach is most appropriate. A

Figure 1
Creativity: Brainstorming

Name: _____			Teacher: J. Aniello		
Date : _____			Title of Work: Uses of a Clothes Pin		
Criteria				Points	
	1	2	3		
Fluency: the number of ideas generated	Less than five	5-10 ideas	More than 10*	_____	
Flexibility: variety of ideas generated	All ideas serve the same basic function	Ideas serve a few (2 or 3) functions	Ideas serve a wide variety of (more than 3) functions*	_____	
Originality: novelty of ideas	Ideas are copies of existing ideas	Ideas are modifications or improvements of existing concepts	Ideas are totally new or even unique	_____	
Elaboration: articulation of ideas	Average person cannot even imagine it	Average person understands the nature of it	Average person can visualize it in his/her mind's "eye"	_____	
				Total →	_____
Comments:					
* This criterion is arbitrary and should be adjusted to fit with the exercise, students' capabilities and faculty's expectations.					

177

classic pre- and post-test using this method (substituting the product in the assignment—for example, using a clothes hanger instead of a clothes pin for the later assessment) administered early and late in the student's program would generate the data necessary to assess this learning objective.

PROCESS:

The creative process is a four-step series of events that happens either consciously or unconsciously. Whether articulated or not, it is generally agreed that these steps are necessary to maximize the creative output of any endeavor. Thus, students' knowledge of the steps of the creative process should help improve their creativity:

Upon completing our program, students will demonstrate their knowledge of the creative process by identifying, defining, and providing examples for each of the four steps in the creative process.

Traditional measures to evaluate knowledge—test questions—could be used to assess this objective.

Knowledge is the first step towards our real goal—application. A learning objective that speaks to application is:

Students, when faced with a business problem, will complete each of the 4 steps in the creative process and derive a viable and creative solution/idea.

This objective could be broken down into the following traits:

Knowledge is the first step towards our real goal—application. The following learning objectives speak to application of the creative process:

1. Without being prompted, students will apply techniques associated with creative processing to discuss and resolve business problems.
2. Students will go through the four steps in a creative process to address a business problem.
3. Students will develop a creative and viable business product idea.
4. Students will articulate and "pitch" an original business product idea to a panel of experts.

A classic manner in which to measure students' ability to optimize creativity through the utilization of the creative process (objectives 2-4) would be to solve a business problem. One that seems to work well is for the creation of a viable new business idea (and plan). Once developed, students can furthermore "pitch" the idea (in person or via memo) to a panel of expert judges for assessment and evaluation.

There are many techniques that facilitate this exercise. I have chosen

to reference the Creative Problem Solving method which Sidney Parnes (1981) built by expanding upon Alex Osborn's (1963) Brainstorming paradigm. The four steps of the process can be described as follows:

1. Preparation: Defined as those preliminary actions necessary from which to build the foundation for creativity. This involves two activities: conducting research (fact finding) and defining the objective (objective finding). In fact finding, we expect the student to find information that will provide a background foundation for the succeeding steps. If one wanted to create a new music system, for example, the existing music systems would first need to be researched. In objective finding, the issue or problem must be identified and priorities set. For example, one goal might be "to increase the music storage capacity of current personal music systems by 500% without increasing the selling price." The criteria used to evaluate performance on this activity as summarized in the rubric below.

2. Incubation: Defined as an unspecified period of time of "Idea Finding." Since creative ideas cannot be forced, this "hatching" period can happen at any time and at any place. It typically occurs when our minds are free from task and when we least expect it, like when showering or jogging. The completion of this stage usually signifies the culmination of the divergence effort, when we have an exhaustive list of possible alternatives. Most experts agree that generating a large number of diverse ideas will optimize the chances for accomplishing a successful outcome to meet the desired objective— much like a photographer knows that taking many pictures increases his chances of getting that one perfect shot. Normally, to produce a photo feature story for National Geographic including about 30 photos, a photographer will shoot 14,000 photos or more. We should not be stingy with our mind's "film" when developing creative ideas—more is better!

3. Illumination: Defined as the "ah-ha" moment when the "Solution Finding" to the problem has been made. It is primarily when the best choice among all alternative ideas generated has been selected. This selection moment is usually the culmination of the convergence effort, whether done individually or as a group. The end result of illumination is the selected idea (in our music storage example it might be an iPod) which, in the next step of the process, will be evaluated according to its ability to meet the stated objectives (500% improvement in storage capacity, same price).

4. Verification: Defined as the stage of "Acceptance Finding" when the idea chosen has proven to solve all of the stated problems and meets all opportunity objectives. It is one thing to feel/think/believe that a solution has been found, but this step is the objective proof of the

hypothesis. In the above example, verification calls for one to build, cost, and test a prototype of the iPod to prove whether it meets the stated storage and price objectives. Thus, at this step, the question is, "Can the idea become a tangible reality?" In many cases, the final step of the creative *Process* (verification) is where the creative *Product* emanates.

The business plan developed by the student could be used to evaluate creative *Process* (see traits 1 through 4). Evaluation criteria for each of these traits are provided in the rubric below.

Figure 2
Creative Process

Name: _____			Teacher: _____		
Date : _____			Title of Work: Business Plan: New Venture		
Criteria					Points
	1	2	3		
Preparation: "fact finding"	No significant research.	At least 3 sources of relevant research.	Research that goes beyond "literal" categories into "lateral" for points of comparison.		_____
Preparation: "problem/ objective finding"	No clear cut problem or objective identified.	A problem identified but no or only vague objectives.	All problems clearly stated and multiple objectives prioritized toward solution opportunities.		_____
Incubation: "idea finding"	Not many ideas generated with little novelty or diversity.	Good number of ideas but not overly novel or diverse.	Many ideas of a diverse nature.		_____
Illumination: "solution finding"	Selection fails to meet objective(s).	Selection solves problem stated in main objective.	Selection meets all objectives/maximizes solutions to all problems.		_____
Verification: "acceptance finding" (idea is proven)	Not a successful solution.	Minimally successful (modification or improvement).	Highly successful solution (uniquely creative).		_____
Product	Not a viable business opportunity.	People willing to invest $$ in the opportunity.	Represents exceptional market potential & exclusivity.		_____
Comments:				Total →	_____

180

PRODUCT:

Much like love being in the "eye of the beholder," a creative product is defined as such by the opinions of acknowledged domain experts in the field of consideration. It is a highly subjective matter analogous to award winners from music, movies, or beauty pageants. If a product (i.e., problem solution) is judged to be creative at this stage, an idea usually begins its life for the journey of innovation that ultimately leads to tangibility. From innovation tangibility, a business opportunity can emerge for an entrepreneur who can recognize it and, hopefully, exploit it profitably.

The business plan project, designed to evaluate *Process*, can be used to evaluate the *Product* dimension of creativity as well. There must be some willingness to invest money in the profit opportunity for the product to meet the objective and be judged "good enough." Product ideas that generate enthusiasm because they represent exceptional market potential would be judged superior. A panel of experts from the local community who are involved in venture capital could add interest to this exercise, in addition to potentially strengthening the external validity of this assessment.

In addition to evaluating student creativity along the lines of three of the 4-Ps, other methods and considerations can also be used as part of creativity measurement standards. For example, a pre-test/post-test can be given as part of an entrance exam/exit exam which asks students to define creativity and explain the importance of creativity as used in a business context. These scores can be then be compared. Additionally, students can be evaluated for their application of creativity techniques during the course of various class projects. This can be assessed through observation and/or creation of a student journal/log as an exercise component of the project.

Closing the Loop

What can a school do if they find that they fall short of their learning goal—that is, students are no more creative in leaving the program than when entering it? The logical first step would be to evaluate the curriculum to see if students receive the necessary knowledge and opportunities to apply and reinforce it. In order for students to understand creativity in the context of the business curriculum, the first thing that they will need is the knowledge of the principles and practices of the topic itself. There are research-proven theories of creativity as well and commonly-accepted methods and techniques which would be beneficial to anyone once exposed to them. It certainly would be helpful to include the study of brainstorming, creative problem solving, lateral thinking or synectics somewhere in the curriculum. This would provide a solid foundation of how and why creativity works.

It remains to be decided whether the above-mentioned learning is best to take place in a dedicated "Creativity 101" course or is better infused throughout all the business courses in the curriculum. Creativity will be a new concept to many students and therefore probably needs to be taught,

and then continually reinforced, over an extended period of time. In some manner, creativity is like the other generalized skills, which are incorporated across the curriculum, such as oral/written communication, ethical reasoning, critical thinking, or information literacy.

In order to anchor or underpin the knowledge of creativity, students will undoubtedly need occasions to apply their knowledge of creativity paradigms. Much like any other talent, students will need to find their creative "voices" in a variety of circumstances and during a number of situations while at school. Coursework tasks, exercises, assignments and projects need to offer ample opportunities for students to exercise their creative "muscles" and practice that which they have learned. This can be done both individually as well as in groups, and probably should be done with both. As mentioned, new business ideas offer a natural chance for students to engage in different, non-linear, thinking patterns.

If a school really values creativity, it is fairly simple to include a "taste" of it in almost any coursework. First of all, there are a number of books which can easily be incorporated into the syllabus. Books like *Corporate Creativity*, by Robinson & Stern; *Embracing Uncertainty* by Clampitt & DeKoch; and *Artful Making* by Austin & Devin would certainly work nicely into the reading list for many business courses. Additionally, there are source books containing creativity exercises that can be well applied to most any discipline for term projects. It is mostly a change of professoriate mindset to begin using creative "games" and other right-brain techniques in order to expand possibilities and generate alternatives to solve business problems.

Once creativity is included into the curriculum, another issue arises. How are these efforts evaluated and graded? We are very clear on how to grade the right/wrong quantitative answers to our questions, but this "grey" area poses an entirely new rubric. It is seemingly more difficult (and time consuming) to assess criteria like uniqueness, synthesis, feasibility, implications, comprehensiveness, and opportunity recognition. Somehow, each faculty member must determine if and how each individual student meets established standards, does not meet standards or exceeds standards. This is a lot more work than giving multiple-choice tests that can be scan-graded. Evaluating creativity necessitates much more individual attention and the management of many more independent work efforts. Adding creativity as a evaluative criterion to assignments across the curriculum—for example, as a component of how a case solution is graded—further reinforces (and rewards) the importance of creative thinking in a business context.

A second important way to "close the loop" with developing students' creative abilities is to develop faculty skills. In order for students to maximize their creativity output, faculty will need to take on expanded roles. Faculty must believe that creativity is important, understand its requirements, model its manifestation and elicit it from students. This may not be a comfortable disposition for many who do not feel creative themselves and may resist

including a non-essential (to them) element into the curriculum. It can be more than a little threatening to give up some control to student autonomy all in the name of creative freedom.

Students' ability to be creative is strongly influenced by their environment—the 4th "P" (*Press*) in the creative discipline. "Environmental factors play a critical role in blocking or aiding the creative process" (Stein, 1974). Individuals need continual outside encouragement to pursue creative endeavors, but society—including education and employment—generally falls short in this regard (Torrance, 1962). Thus, the faculty member's role goes beyond creating a curriculum that teaches students about creative process and techniques, developing exercises and assignments to give students the opportunities to apply this knowledge to creative problem solving, and incorporating creativity into how students' work is evaluated. Creating a classroom environment that is supportive of creativity is, arguably, the most critical thing a faculty member can do to develop a student's creativity.

Faculty influenced conditions most contributing to a positive creative environment are exposing students to a variety of experiences, encouraging creative activities, collaborating with students, supporting efforts with appropriate resources (most important is access to the faculty member), respecting individuals, exhibiting confidence in students' abilities, and recognizing and rewarding creativity (MacKinnon, 1966). Some of these are issues involving curriculum design (designing experiences, activities and, to some extent, collaboration), but the final three preconditions listed are a direct offshoot of the faculty member's interaction with the student, and the key component of this must be *trust*. Drawing on concepts developed by Carl Rogers (1959), in order for individuals to take the risks and experience the discomfort of tapping into their latent creativity, their environment must allow for *psychological freedom* and *psychological safety*. Psychological freedom exists when all creative expression is welcome, while psychological safety exists when individuals feel understood and valued, and can be nurtured without the fear of external evaluation.

How can faculty members create such environments in their classrooms? Carkhuff's "Helping Hand" model (Carkhuff, 1981) builds upon Rogers to develop some specific, common sense recommendations on how to improve the faculty-student interaction that is directly relevant to creativity. The Helping Hand model calls for the faculty member to create an environment that is natural and comfortable for the student, where instruction is delivered in small steps, builds on students' comprehension, and transfers much of the faculty member's knowledge to students. Finally, learning must culminate in the successful achievement of an agreed upon objective (Carkhuff, Berenson & Pierce, 1976)—approximating the transition from creativity to innovation in the business environment.

In creating a supportive environment, faculty members not only encourage (or require!) that students engage in exploration of their own

creativity, but they can influence student motivation and interest to be creative. Since creativity is an ongoing process, motivation is, by far, the best predictor of success in this area, outpacing knowledge or even creative talent (Amabile, 1996). By providing students with successful creative experiences, faculty help develop their students' interests and confidence in their own creative abilities, thereby improving the students' internal drive to act upon them and be creative. Sufficient internal motivation, combined with a supportive environment, can result in "passion" for one's own activity—that characteristic that defines success not only for entrepreneurs, but in so many fields.

Thus, if creativity is to be encouraged in the curriculum, the appropriate classroom environment is necessary. In providing the necessary psychological support and social infrastructure, along with instruction and feedback to students as they engage in the creative process, faculty must provide the crucial link to developing students' creative potential. Creativity is not so much taught as it is encouraged and supported by faculty who do so by their example. Faculty, at their best, can be talent developers who inspire students to use a multiplicity of their abilities and allow creativity to be the process in which knowledge unfolds.

This is a daunting challenge for most of us but, due to the pivotal role of the faculty member in the creative process, it must be met. Faculty development can play a very important role in "closing the loop" to improve the creativity curriculum. Two organizations that can help immensely in faculty development are the Creative Education Foundation and the Center for Creative Leadership.[52] These organizations' Web sites include links to a wealth of resources available for faculty development in the area of leading and facilitating creativity from book libraries, to seminars, to research studies, to key contacts in the field. If truly dedicated to expanding creativity within the curriculum, higher education administrations will commit more time, space and funds to meet the challenge of faculty development.

Conclusion

In conclusion, creativity—which will be the key ingredient to the success of all types of organizations in the future—deserves a place in the business curriculum. In schools with an entrepreneurial mission, creativity should have a prominent place, be incorporated across the curriculum, and be among the school's learning goals for its students. Creativity should not be confused with artistic talent—everyone has the necessary creativity in order to develop better solution alternatives to increasingly more complex problems. Creativity can be improved and encouraged and, since most people have not realized their potential, there is great potential for improving students' creative abilities. A combination of curriculum design and faculty development can result in a degree program that produces more creative graduates. Finally, as I hope the examples in this chapter demonstrate, learning objectives associated with creativity are not difficult to assess and can provide valuable feedback to

184

stakeholders (including potential employers) that the school has delivered on its creative or entrepreneurial mission with respect to student learning.

Endnote

52 To access more information on the Creative Education Foundation go to http://www.creativeeducationfoundation.org. The Center for Creative Leadership can be found at http://www.ccl.org

References

Amabile, T. M. (1996). *Creativity in context.* Boulder, CO: Westview Press.

Barron, F. (1969). *Creative person and creative process.* New York: Holt, Rinehart & Winston.

Carkhuff, R. R. (1981). *Toward actualizing human potential.* Amherst, MA: Human Resource Development Press.

Carkhuff, R. R., Berensen, D. H., & Pierce, R. M. (1976). *The skills of teaching.* Amherst, MA: Human Resource Development Press.

Isaksen, S. G. (1988). Innovative problem solving in groups: New methods and research opportunities. In Y. Ijiri & R. L. Kuhn (Eds.), *New directions in creative and innovative management: Bridging theory and practice* (n.p.). Cambridge, MA: Ballinger Publishing.

Lynch, D., & Kordis, P. L. (1988). *Strategy of the dolphin.* New York: William Morrow.

MacKinnon, D. W. (1966). Instructional media in the nurturing of creativity. In C. W. Taylor & F. E. Williams (Eds.), *Instructional media and creativity* (n.p.). New York: John Wiley & Sons.

Osborn, A. F. (1963). *Applied imagination* (3rd ed.). New York: Charles Scribner's Sons.

Parnes, S. J. (1981). Creative problem solving. In S. S. Gryskiewicz & J. T. Shields & W. H. Drath (Eds.), *Selected readings in creativity from the first five years of creativity week (Vol. II)* (n.p.). Greensboro, NC: Center for Creative Leadership Press.

Rogers, C. R. (1959). Toward a theory of creativity. In H. H. Anderson (Ed.), *Creativity and its cultivation* (n.p.). New York: Harper & Brothers.

Runco, M. A., & Sakamoto, S. O. (1999). Experimental studies of creativity. In R. J. Sternberg (Ed.), *Handbook of creativity* (n.p.). Cambridge, UK: Cambridge University Press.

Stein, M. I. (1974). *Stimulating creativity in individuals.* New York: Academic Press.

Stewart, T. A. (1997). *Intellectual capital.* New York: Doubleday.

Torrance, E. P. (1962). *Guiding creative talent.* Englewood Cliffs, NJ: Prentice-Hall.

Torrance, E. P. (1974). Creative teaching makes a difference. In J. C. Gowan & J. Khatena & E. P. Torrance (Eds.), *Creativity: Its educational implications* (n.p.). Dubuque, IO: Kendall/Hunt.

Wheatley, M. J., & Kellner-Rogers, M. (1996). *A simpler way.* San Francisco: Berrett-Koehler.

Author Bio

Dr. Joseph Aniello is Assistant Professor of Management at Francis Marion University, where he teaches Organizational Behavior and Entrepreneurship and is Director of the Donald E. Kelly Small Business Institute. Prior to entering academia, Dr. Aniello held management positions at Gibson Guitars, Avon Products, Saatchi & Saatchi Advertising and other corporations.