

ISSUES IN INTEGRATIVE STUDIES
No. 20, pp. 77-101 (2002)

Interdisciplinary Program Assessment

by
Donald E. Stowe

University of South Carolina
College of Hospitality, Retail, and Sport Management

and

Douglas J. Eder

Southern Illinois University, Edwardsville
Undergraduate Assessment and Program Review

Abstract: Interdisciplinary studies programs and the assessment movement are two parallel educational paradigms, both of which are maturing at a pace that is noticeably deliberate. Idealistic in inception, they are simple in concept, nuanced in practice, and focused on student learning rather than on professorial teaching. Idealism and subtle complexities in interdisciplinary thinking contribute to a belief that assessment is too simplistic to court interdisciplinary studies and has little business here. This paper questions the veracity of that belief and offers some research-based assessment approaches that may sanction an acceptable, if not lusty, marriage of the two paradigms.

Introduction

AS THE ASSESSMENT MOVEMENT in the United States enters its fourth decade, countless hours and numerous pages continue to champion the as yet unfulfilled promise of improving teaching and learning through assessment. Although the vigorous adoption of an assessment ethic by regional accrediting agencies and governing boards has certainly reified the process, educational organizations can still pay obedient lip service to assessment and get away with that and precious little else. A survey by the National Center for Postsecondary Improvement (1999) of 1300+ chief academic officers revealed a consistent opinion that the assessment movement has had little impact on improving student learning. When assessment data of any character did exist, very little use has been made of the data for teaching and curriculum

improvement.

While the waves of educational reform have ebbed and flowed over the past thirty years, three ideas—the movement from teaching-based to learning-based education, assessment, and the interdisciplinary studies (IDS) movement—have together weathered an extended adolescence. The parallels of these reform ideas are remarkable. They are the progeny of idealists who seek to make education more responsive and more authentic. However, after much discussion and public acclaim, and while the three ideas weigh on the minds of many, actual implementation varies widely. Despite the reticence to move from thought to action, the potential integration of interdisciplinary education, an outcomes model of learning, and assessment presents an exciting opportunity to improve teaching and learning.

Concurrent with the work of others who have championed a paradigm shift from teaching to learning or the implementation of assessment, the Association for Integrative Studies (AIS) has worked diligently for the past twenty years to put a public face on the otherwise diffuse idea of interdisciplinarity. Accomplishment of this ideal is a monumental task—one worthy of the honest time, effort, and scholarship required to brighten that face. In late fall of 1998, AIS formed a task force to determine the status of assessment within the interdisciplinary community and to offer suggestions for enhancing the appropriate use of assessment within that community.

This paper is the result, to date, of three and one-half years of ongoing consideration of assessment by the AIS Assessment Committee. Having immersed ourselves in the parallel ideals, rhetoric, and realities of assessment and of interdisciplinarity, we have come to appreciate the complexity and diversity of opinions and options that exist. Initial inquiries by the AIS Assessment Committee in 1999 revealed that assessment within the interdisciplinary community was in a protracted infancy and that few mature assessment plans existed. Many interdisciplinary programs were “caught up” in some sort of external mandate to implement an assessment strategy, and these programs were eager for thoughtful advice. While the interdisciplinary movement must not be hampered by overly prescriptive or overly simplified ideas regarding assessment and learning, the potential to meld the needs of interdisciplinarians struggling to achieve a mainstream position in academia with the ethos of assessment is compelling. Yet, actual implementation of a viable assessment plan remains a challenge to many of our colleagues.

There are a number of reasons that interdisciplinarians find it difficult to mount a sustained assessment effort. Foremost among them is that the construct of interdisciplinarity is actively disputed. Interdisciplinarity is a

complex idea, one that is not readily subjected to the seemingly reductive processes of assessment (Klein 2002). This, coupled with the reality that some faculties simply don't want to do assessment, presents a formidable challenge. Other problems stem from the local nature of interdisciplinarity. Stakeholders who casually adopt the rhetoric of interdisciplinarity often have not constructed their own mission well enough to articulate it in the detail required to develop an effective assessment plan. It is fascinating to get a group of well-meaning interdisciplinary faculty members in the same room and launch them on a discussion of their mission. As they speak wisely and respond knowingly to the comments of their colleagues, it often becomes apparent that individuals are really talking about their own vision of the mission and that as a group they have never really reached consensus. The mission of an IDS program is a local matter that must be defined, not through some absolute reality, but through the process of negotiation and consensus.

Resolving Some Basic Issues

Before we lay more specific groundwork for assessing interdisciplinary programs, we will identify and discuss certain issues that are constantly on the forefront of assessment, generally in higher education. We do this because many IDS programs, when faced with the challenge of developing and implementing an assessment plan, become mired in the rhetoric of assessment and interdisciplinarity and fail to resolve more fundamental issues. Consequently, they make the most common error in assessment—the failure to get started. Issues that may cause an assessment plan to die before it starts are:

- What is assessment?
- What is the difference between assessment and evaluation?
- Must measurement be the single perspective of assessment?
- Do we aim our assessment plan primarily at the course, program, or institutional level?

We will address these issues one at a time.

What Is Assessment?

While definitions abound, it must be clear from the outset that the definition of assessment used on one campus may not work well on another campus (Palomba & Banta 1999, p. 3). We hasten to add that definitions of assessment also may vary from one interdisciplinary program to another. Nevertheless, certain definitions seem to correspond nicely with the ambitions of

the interdisciplinary community.

The idea of assessment took an evolutionary turn in the spring of 1995, when Tom Angelo, then director of the American Association for Higher Education (AAHE) Assessment Forum, suggested it was time to “reassess assessment in higher education” (1995a, p. 11). Angelo presented the following draft definition of assessment to the readership of the *AAHE Bulletin*. Colleagues were invited to comment on, revise, and/or expand the definition.

Assessment is a means for focusing our collective attention, examining our assumptions, and creating a shared culture dedicated to continuously improving the quality of higher learning. Assessment requires making expectations and standards for quality explicit and public; systematically gathering evidence on how well performance matches those expectations and standards; analyzing and interpreting the evidence; and using the resulting information to document, explain, and improve performance. (p. 11)

As promised, Angelo took seriously the feedback he received. Five themes emerged from the contributions:

1. Assessment should focus on improving student learning.
2. The focus of assessment should not be limited to the classroom, but include the wide range of processes that influence learning.
3. Assessment is a process embedded within larger systems.
4. Assessment should focus collective attention and create linkages and enhance coherence within and across the curriculum.
5. Tension between assessment for improvement and assessment for accountability must be managed.

Angelo’s revised definition reads as follows:

Assessment is an ongoing process aimed at understanding and improving student learning. It involves making our expectations explicit and public; setting appropriate criteria and high standards for learning quality; systematically gathering, analyzing, and interpreting evidence to determine how well performance matches those expectations and standards; and using the resulting information to document, explain, and improve performance. When it is embed-

ded effectively within larger institutional systems, assessment can help us focus our collective attention, examine our assumptions, and create a shared academic culture dedicated to assuring and improving the quality of higher education. (1995b, p. 7)

We recommend Angelo's extended discussion to those who are struggling with the idea of assessment and/or preparing to communicate the essence of assessment to a novice audience. For those who enjoy some experience with assessment and some local acceptance of the idea, a shorter definition may serve as well. For example, Marchese suggested that assessment is "the systematic collection, review, and use of information about educational programs undertaken for the purpose of improving student learning and development" (qtd. in Palomba & Banta 1999, p. 4). Whether one is a novice or experienced in the assessment arena, the discussion surrounding the construct of assessment is ongoing. Regardless of the local conception of assessment, the AIS Assessment Committee commends working definitions of assessment that emphasize improving student learning.

What Is the Difference Between Assessment and Evaluation?

Those who casually observe the processes of assessment often confuse them with evaluation. Observers can be frustrated by the apparent similarities while hearing the claim that, in fact, assessment and evaluation are entirely different. Indeed, there are similarities between the two, and data collected for one purpose can potentially be used for the other. In general, when data are collected to improve teaching and learning, the process is assessment. For our purposes, this data collection can range from a single professor engaging in classroom assessment to an entire institution undertaking a self-study, and the information may be collected at various points in the teaching-learning process. Because adjustments of pedagogy or curriculum can be made along the way, this is often referred to as a formative process. The motivation for assessment should be pure and uncorrupted by threat of external judgment; else, one loses the ability to ask honest questions.

In contrast, when data are collected to make judgments of worth or value, the process is evaluation. Judgments might involve qualification for tenure, continuation of a program, the level of funding, or continuation of accreditation. Data for the purpose of evaluation are collected in a manner that summarizes an entity's qualities. If the name "assessment" is applied to

such a process, it is called *summative assessment*. While it is clear to us that the same data can be used for both summative and *formative* processes, differences lie in the nuances of the processes and the outcomes.

Must Measurement Be the Single Perspective for Assessment?

Some people entering the assessment arena for the first time have previously had conventional measurement training. While adherence to the precepts of psychometrics is responsible and even admirable, the diversity of IDS programs makes unexamined adherence to measurement principles unwise. Later in this document we will present five perspectives that one may employ in the process of assessment—measurement among them. However, many assessment programs have been crippled or doomed from the start because of a blinkered adherence to quantitative measurement. The AIS Assessment Committee recommends that IDS programs consider a variety of perspectives that exist along a continuum of quantitative and qualitative assessments. Those perspectives will be detailed later in this paper.

Do We Develop Our Assessment Plan at the Course, the Program, or the Institutional Level?

A coalescence of opinion in several venues, for instance, the Association of American Colleges and Universities' report (2002) entitled, "*Greater Expectations: A New Vision for Learning as a Nation Goes to College*," offers the view that what faculty members want to see in students is evidence of "deep learning." This kind of learning is relational as well as absolute. Yes, it requires mastery of facts, yet it relates those facts to other facts within and without the discipline. Deep learning invokes facts when necessary, sorts them, and recognizes when other necessary facts are absent. Persons who possess deep learning also know how to go after those missing facts. In short, deep learning is the learning possessed by experts in a field, and it is usually acquired directly through instruction and indirectly through experience. If deep learning is what faculty members seek, then several factors need to be present in the collegiate environment including, but not limited to, these:

1. Students must experience a pedagogy beyond "standard" passive lecture.
2. Students must face a significant problem for which their own discipline and its way of knowing is a necessary, but not suffi-

cient, contributor to the solution.

3. Students must see the need for acquiring relevant and important facts outside their own major disciplines and inserting them into new contexts.
4. Students must practice acquiring those facts and manipulating them in those contexts under faculty supervision.

How might the faculty in an institution of higher education know whether its students have attained “deep learning”? One response from the perspective of assessment affirms that deep, expert learning is a system function, not a course function. Colleges and universities are systems, and the total education received by students is a function of the whole learning environment, both inside and outside the classrooms. The well-known Annenberg Foundation short film called *A Private Universe* (Schneps & Sadler 1987) reveals humorously and poignantly how the facts learned in single courses peel off easily like veneer in the rain. Once a course is finished, even bright students who scored A’s commonly revert to old, comfortable prejudices and superstitions. Hence, if deep learning is really what we want, then assessment of deep learning properly focuses at the departmental, program, or institutional level rather than on individual courses. Almost by definition, interdisciplinary learning is deep learning, so IDS assessment may best be done at levels above individual courses.

Specific Recommendations for Interdisciplinary Assessment

Universities have historically arranged themselves into units that reflect academic disciplines, their attendant cultures, and their specialized ways of knowing. The resulting department-based structures serve the disciplines well, but they do not serve well when attempting to work across the disciplines. Thus, effective assessment of interdisciplinary studies has lagged behind assessment in the disciplines. Moreover, some cross-cutting concepts embedded especially in interdisciplinary studies, such as life-long learning, curiosity, creative thinking, synthesis, and integration, have acquired an aura of being ineffable and, correspondingly, unassessable. It seems to us, however, that assessing interdisciplinary studies is not so much a challenge of assessing the ineffable as assessing the inchoate. Ineffable means “indescribable,” whereas the word inchoate comes from the Latin *cohūm*, meaning a strap used to hitch an item like a plow to the ox that pulls it. Therefore, inchoate means “not hitched up.” It is becoming increasingly clear that the maturing

field of interdisciplinary studies can, with increasing clarity, be described. Assessment principles just haven't been "hitched up" to IDS yet, a process that awaits the required imagination, energy, and perseverance.

As we appreciate mainstream assessment and begin the long process of developing a specific scholarship of interdisciplinary assessment, three technical ideas arise that may benefit by interpretation from an interdisciplinary perspective: approaches to data collection, instruments/techniques of assessment, and models for assessment. We discuss these issues, not to be overly prescriptive, but to offer insights that may make assessment more palatable to the interdisciplinary community.

Approaches to Data Collection

One of the larger misconceptions still arising during discussions of assessment is that there is a singular approach to appreciating data that transcends all other perspectives. The role of conventional measurement theory has provoked an ongoing discourse (Biggs 2002, National Research Council 2001). Farmer and Napieralski (1997) identified several perspectives on assessment that we will modify for the purpose of looking at the assessment of interdisciplinary programs (p. 593). Five perspectives that one might consider are Assessment by Measurement, Assessment by Objectives, Assessment by Professional Judgment, Assessment by Questions, and Assessment by Discovery. These perspectives, though not entirely linear or continuous, range roughly along a continuum: from scientific to humanistic, from quantitative to qualitative, from objective to subjective (see Table 1).

Table 1

Assessment by Measurement	Assessment by Objectives	Assessment by Professional Judgment	Assessment by Question (Descriptive)	Assessment by Discovery (Goal-Free)
Scientific ←————→ Humanistic				
Objective ←————→ Subjective				
Linear ←————→ Non-linear				
Quantitative ←————→ Qualitative				

Each of these approaches has merit in the process of assessing interdisciplinary programs that by their nature have ambitious and complex outcomes.

Each approach has different assumptions, constraints, and anticipated results (1997, p. 593). A well-conceived, continuously evolving IDS assessment program will possess, at one point or another, some aspects of each approach, with the emphasis shifting as needed. Discussion of these five perspectives concludes with a summary, found below in Table 2.

Assessment by Measurement assumes that the program has measurable learning outcomes and that instruments can be designed or may already exist to measure the outcomes in a reliable and valid manner. Although strengths abound for this approach, the emphasis on technical issues may discourage some faculty members and students from participating in the process. Thus, discussion may be relegated to external stakeholders who place more emphasis on accountability than on improving teaching and learning. Also, the ready presence of refined instruments may have the impact of causing programs to default their intended learning outcomes to those constructs measured by existing instruments rather than struggling to gain data on desired outcomes that are important but appear difficult to measure.

Assessment by Objectives is closely akin in spirit to both the measurement approach and the assessment movement's avowed intention to gather data on intended learning outcomes and to use that data to enhance learning. Application of this approach will require that IDS faculty, students, and other affected stakeholders become involved in defining intended learning outcomes and developing instruments, especially rubrics and protocols, to gather data. The standard for data collection may be less technically rigorous than the standards of the measurement approach. The move from nationally normed instruments to locally developed instruments is prominent in this approach.

Assessment by Professional Judgment, or what Eisner (1998) calls the *Connoisseurship Model*, emphasizes experience, judgment, and intuition over measurement concerns. Interestingly, the word *assess* means to "assist the judge" or "sit beside." Yet we are often inclined to let the data replace rather than assist the judge. Assessment by Professional Judgment moves out of the domain of psychometrics and statistics and into a realm where the observer is indeed a judge—a skilled questioner—rather than an instrument of repetition. Although the role of professional judgment is highly valued in higher education and reified in the form of the "expert consultant," some are reluctant to apply reasoned judgment to the assessment process, ostensibly because of its apparently subjective nature. However, we value expert judgment in other phases of our lives without question: umpires call balls and strikes, judges render informed opinions in gymnastics, the judiciary renders opinions on complex matters of law. It is unusual that we would seek to rid

an assessment system of human judgment for the sake of statistical perfection (Wiggins 1998, p. 111). While the lack of objective criteria may inhibit many, this approach may be effectively applied in concert with reflective scholarship and more objective techniques to provide insights and depth not readily available with one perspective or technique alone. Gordon Vars (2002) has thought extensively about the application of this measurement perspective to the assessment of that daunting IDS outcome—synthesis. He cautions, though, that a concept as complex as synthesis probably cannot be assessed with a high degree of precision. The more complex the construct, the less precise the measurement is likely to be. Vars argues persuasively for the artistry of teaching and of assessment (2002).

As one moves further along the continuum, a decidedly dualistic view of assessment emerges. One is drawn to the possibility of assessing intended outcomes while at the same time allowing for the emergence of unintended outcomes. Goal-Free Assessment allows for the identification and assessment of outcomes other than those strictly expected. A dualistic view of assessment, which acknowledges the assessment of intended outcomes even while allowing for the emergence of unintended outcomes, is entirely consistent with the ambitions of many interdisciplinary programs, and it forms the basis for the following views on data collection.

Assessment by Question, or what Schilling and Schilling (1998) call *descriptive assessment*, also moves father away from the rigors of measurement and provokes a highly creative environment in which the intended outcomes of assessment may be pursued. Unintended outcomes may be discovered in this manner as well. The platform for assessment here is a question or questions that stakeholders find interesting or in some cases perplexing. Questions may range from “What have our students learned this year?” and “What are our students’ metaphors for their educational experiences with us?” to more focused questions such as “Are there differences between the learning outcomes of our residential and commuter students?” and “What is the ‘job description’ for our first year students?”

Assessment by Discovery, in many ways related to Assessment by Questions, allows a program to acknowledge that learning outcomes are not limited to those anticipated or intended. Often called Goal-Free Assessment, discovery allows one to assess intended as well as serendipitous learning outcomes. Goal-free Assessment has been around for some time, having been documented as a companion to Tyler’s (1949) goals-based evaluation by Scriven (1974). While a strict interpretation of Goal-Free Assessment allows one to witness and thereby assess intended outcomes of a program, Assess-

ment by Discovery primarily encourages one to roam among collected data without a significant expectation of what may be found.

Table 2. Summary of five perspectives on assessment

	Assessment By Measurement	Assessment By Objectives	Assessment By Professional Judgment	Assessment By Question (Descriptive)	Assessment By Discovery (Goal-Free)
Attributes	-Formal -Systematic -Emphasis on reliability, validity, and norms -Analytic -Emphasizes data and instruments	-Focus on intended outcomes -Analytic	-Emphasis on experience and judgment -Holistic	-Focus on questions of importance to faculty and students	-Goal-free
Strengths	-Objectivity -Availability of instruments	-Stakeholders involved in reaching consensus on intended outcomes before instruction begins	-Offers access to complex outcomes	-Offers flexibility to stakeholders who desire assessment and are not comfortable with measurement	-Acknowledges potential for unexpected outcomes -Validates intended outcomes
Weaknesses	-Focus on technical concerns/process may negate the purpose of assessment -Intimidating to those not trained in measurement	-Difficulty creating measurable outcomes for complex intellectual tasks	-Subjectivity -Observer bias	-Subjectivity -Observer bias	-Subjectivity -Observer bias
Data Collection	-Standardized instruments—both direct and indirect	-Rubrics -Performance criteria -Observational protocols	-Informed opinion -Rubrics -Performance criteria	-Qualitative -May be informed by quantitative data	-Qualitative

Environments for Assessment: Direct and Indirect Methods

With some understanding of different perspectives on assessment, we move to a more specific discussion of data collection. A frequent phrase in assessment literature refers to activities ranging from standardized tests, surveys, questionnaires, focus groups, locally developed instruments, portfolios, and performances as assessment techniques. It is probably wise at this point to distinguish between two large classes of assessment techniques, namely those that are direct from those that are indirect. *Direct assessments* acquire evidence about student learning and the learning environment itself. Examples

are exams, projects, logs, portfolios, and direct observations. *Indirect assessments* acquire evidence about how students feel about learning and their learning environment. Examples include surveys, questionnaires, interviews, focus groups, and reflective essays. For the most part, colleges and universities invoke indirect assessments in support of their statements about student learning. Accordingly, individual professors hear the word assessment as meaning indirect assessment, that is, mostly surveys and questionnaires, which are undertakings by non-faculty offices, such as Institutional Research. Even when direct assessments, such as standardized tests, are used, they usually do not involve professors as direct participants. Consequently, feedback from assessment efforts is commonly either not received by faculty members or is ignored cheerfully.

While certain standardized tests and surveys developed specifically for the purpose of assessment may be thought of as techniques, the range of environments in which one may gather data extends far beyond tests and surveys. We take the position that multiple sources of data must be collected in interdisciplinary assessment (tests and surveys among them, of course), and that environments created naturally as a consequence of interdisciplinary instruction may already contain the data needed to assess interdisciplinary learning. These environments, which may be examined through one or more of the five assessment perspectives above, include both the direct and indirect techniques—and many more. Most importantly, the feedback acquired from such assessments must surface in a timely and visible manner and become available to those who can benefit from it.

Models for Assessment

In addition to an appreciation of the range of data collection perspectives and environments that may help improve interdisciplinary learning, one must possess a plausible model that will allow for an exploration of the complex relationships that exist among curriculum, pedagogy, and learning. One of the first rules of assessment is to start simply and develop a more sophisticated and evolved plan as practice allows. The following discussion assumes that an initial interdisciplinary assessment plan may begin with the gathering of data on a few learning outcomes deemed important locally. To wit, one might ask, “What three (or five, or eight) things do we wish our graduating students to retain from our program five years from now?” In other words, what major learning outcomes do we want students in our program to achieve? However, as local experience with assessment evolves, it is inevitable that the players will begin to ask appropriate questions about the intricate rela-

tionships among student characteristics, curriculum, pedagogy, and environment. The work of Alexander Astin (1991) and his colleagues offers a model, *the outcomes-only model*, that enhances this process and may promote assessment of complex interdisciplinary programs without resorting to oversimplification or undue complexity. An initial awareness of the potential interactions among the major components of the model bodes well for future enhancements of a local assessment plan. Every plan must have a starting place, and we strongly recommend beginning with such an outcomes model of assessment.

Astin provided the academy with a model that interacts well with the assessment definitions of Angelo and others and seems to respond to the current needs of interdisciplinarians interested in assessment. Astin's model postulates three interrelated ideas: INPUT-ENVIRONMENT-OUTCOMES or I-E-O. Astin contends that this model provides a "powerful framework for the design of assessment activities and for dealing with even the most complex and sophisticated issues in assessment and evaluation" (p. 16). The premise of Astin's model is that learning outcomes must be understood in terms of the characteristics of the learner and the elements of the environment in which the learning occurs, especially curriculum and pedagogy.

Inputs may include gender, race, age, SAT or ACT scores, family socio-economic status, high-school grades, aspirations, and personal motivation. Environment is that complex area where philosophy, content, curriculum, and pedagogy occur. Since we assume that there are some relationships among curriculum, teaching, and learning, the complexity of the environment must not be ignored in the assessment process. Outcome refers to intended and unintended results and may be classified at four levels: cultural, institutional, programmatic, and course. Cultural outcomes are important to acknowledge and difficult, but not impossible, to assess. They include ideals such as citizenship, confidence, and character. Institutional outcomes must be conceived in concert with local mission and may be as general as retention and graduation or as specific as the institutional conception of general education. Outcomes at the programmatic level are the primary focus of our efforts here, for they focus explicitly on learning outcomes. In the interdisciplinary domain, learning outcomes may range from conventional learning, such as written and mathematical competencies, to the ultimate outcome of an interdisciplinary education—synthesis or integration. In some circles of the higher education community there is an emphasis on assessment at the course level. By no means do we wish to discourage this possibility, but we point out that most of the intended outcomes of interdis-

ciplinary programs are laudable and complex. It is unlikely that an emphasis on course-by-course assessment will be as effective in improving learning as will an emphasis at the program level.

Astin argues that teaching and learning cannot be continuously improved by the application of parts of the model. Given the diversity and complexity of interdisciplinary programs and intended/suspected outcomes, the Astin model has merit for the interdisciplinary community. Astin defines commonly found but incomplete assessment designs as:

1. Outcomes-Only Assessment
2. Environment-Outcomes Assessment
3. Input-Outcomes Assessment
4. Environment-Only Assessment

The Outcomes-Only model is widely touted in assessment. Ultimately, the AIS Assessment Committee is inclined toward models that examine the intricate relationships among student characteristics, curriculum, pedagogy, and outcomes. However, it is more than fair to say that the art of assessment is in its infancy, especially in the area of interdisciplinary studies. Rigid insistence on implementation of the complete model is likely to stall the development of interdisciplinary assessment plans. Thus, most will want to work first with the outcomes component and then proceed later to more sophisticated applications of available models.

The Interdisciplinary Assessment Plan

Up to now we have sought to define terms; quiet some troublesome issues; and put an interdisciplinary slant on the areas of data collection, assessment techniques, and models of assessment. Conventional assessment wisdom mandates that an assessment plan be developed before any activities are actually undertaken. The assessment plan must be a living document; data and discourse must be in constant motion. As the plan evolves, the notion of continuous quality improvement becomes a reality.

There are many frameworks for assessment, each having merit. In our earlier discussion we acknowledged assessment perspectives that focus, at the extremes, either on questions of interest to the faculty (Descriptive Assessment) or on unknown outcomes (Goal-Free Assessment). One of the paradoxes of interdisciplinary assessment is that it may be beneficial to focus in the early stages of assessment on an approach that is more linear and more outcomes oriented, despite our acknowledgment of the complexities of

IDS. An early focus on mission and anticipated outcomes may cause a faculty to come to grips with some unspoken assumptions and cause professors to work toward a degree of consensus on outcomes. For those who conceive their mission thus, and prefer a more non-linear approach, the notions of Descriptive and Goal-Free Assessment are recommended. However, the discussion that follows will have a decidedly linear bias as it contemplates some of the most important notions in assessment and the unique circumstances posed by interdisciplinary programs. This paper will proceed based upon the following questions that each interdisciplinary program should answer as it lays the foundation for making assessment an integral part of the teaching-learning paradigm.

- What is the mission of the IDS program? For whom and why does the program exist?
- What are the major goals of the IDS program? What should a student be able to know, feel, think, and do after completing the program?
- What are the specific objectives of the program? What are the measurable indications that the goals are being achieved?
- What data will we collect? How will we collect data to see if the goals/objectives are being achieved?
- When will we collect data?
- What will we do with the data?

What Is the Mission of the Interdisciplinary Program?

Mission statements in higher education are notoriously vague (Dickeson 1999). While the virtues of vagueness are limitless, the liabilities in terms of IDS assessment are abundant. A program unable to articulate, at the minimum, whom it intends to serve, how those persons will be served, and what one may generally expect to happen as a consequence of being served, is unlikely to get past the first stages of developing a useful assessment plan. Spending time in committee or on a retreat answering the following questions is an excellent faculty development process that will lay a foundation for an effective interdisciplinary assessment plan.

Whom Will the Program Serve?

Interdisciplinary programs serve diverse populations, i.e., first and second year students, adult learners, honor students. The list is endless. Since the Astin model emphasizes input, it is important to consider the characteristics

of the students being served. Demographic variables may be collected if the program considers them important.

How Will the Students Be Served?

This question speaks to issues of curriculum and pedagogy. A clear answer to this question, based on local consensus, is critical. It forces the program faculty to focus upon a local conception of interdisciplinarity and to make explicit the learning outcomes that they value and expect. The juxtaposition of interdisciplinary and assessment theory in the context of this question can provide an unmatched impetus to faculty and program development.

What Generally Do We Wish Students to Learn?

This section of the mission statement does not have to be an all-encompassing statement of learning outcomes. This is the section of a plan that outlines the intended general goals of the program, and it should make a good-faith attempt at putting the learning outcomes of the program before the public. The mission statement may be viewed as an executive summary, ever popular in higher education today. A concise statement about the intended outcomes of the program is another step along the way to developing an effective plan.

What Are the Major Learning Goals of the IDS Program?

One of the barriers to effective assessment among interdisciplinary programs is the failure to distinguish between potential levels of assessment. Institutions may have assessment plans at several levels: institutional, program, course, classroom, and individual assignment or student. Assessment at each of these levels can provide information that informs the other levels (Palomba & Banta 1999, p. 9). Although goals for each level are important, this paper will focus on learning outcomes at the program level, given the proximity of program learning outcomes to our accepted assessment goal of improving teaching and learning.

Learning outcomes at the institutional level are heavily influenced by local mission, but most include ideals such as the development of citizenship, character, and confidence. Learning outcomes at the program level may include more specifically defined cognitive, affective, and developmental outcomes. One of us (DJE) has observed that among nearly 80 institutions of higher learning, only a few mention affective or developmental aspects in their assessment plans at all, and most of these have an extensive assessment

history (e.g., Alverno College), a vital school of education, or an influential program in interdisciplinary studies. While learning goals may also be defined at the assignment, course, classroom, and individual levels, a focus on assessment at this level, while valuable, is not the specific intent of our efforts at this time.

What Are the Specific Learning Outcomes of the IDS Program?

Newell (1998) acknowledges that the outcomes of interdisciplinarity are numerous, diverse, and subject more to claim than to proof (p. 537). Learning outcomes that do occur as a consequence of interdisciplinary study are certainly not the unique or sole province of interdisciplinarity; they may occur as a consequence of traditional disciplinary learning as well. However, interdisciplinarians argue that interdisciplinarity is a more authentic way of conceptualizing knowledge, and that the potential learning outcomes are likely to be richer, more complex outcomes than those achieved as a consequence of disciplinary study.

Broadly speaking, the learning outcomes of IDS can be categorized as cognitive, affective, and developmental. Potential cognitive outcomes of IDS include, but are not limited to critical thinking, problem solving, contextual understanding, coping with complexity, making connections among discipline-based ideas, creative thinking, facility with analogy and metaphor, awareness of biases, sensitivity to moral and ethical issues, awareness of values, and the ability to synthesize or integrate disciplinary insights to construct a more comprehensive perspective (Newell 1998). Squires argued that the litmus test of interdisciplinarity is the presence and extent of integration of disciplinary perspectives (1992, p. 206). The general acceptance by mainstream interdisciplinarians of Squires' contention positions synthesis as a priority in any IDS assessment plan. In addition to the highly valued cognitive outcomes of IDS mentioned previously, assessors should not ignore more pragmatic yet essential outcomes such as writing, speaking, computing, and using the scientific method. Potential affective outcomes of IDS include increased self-confidence, a sense of empowerment, respect for differences, and ability to work on a team. Developmental outcomes may include moral and ethical development and progression through a series of developmental stages to a theoretically defined apex (Baxter-Magolda 2001).

Interdisciplinary programs may have a number of agreed upon cognitive, affective, and developmental outcomes at the level of student learn-

ing. The outcome most central and fundamental to mainstream IDS programs is the notion of synthesis/integration. Synthesis may be the most discussed and least-well documented term in the interdisciplinary literature (Newell 1998). Everything we look at is merely a proxy for the desired learning outcome. It is possible that synthesis is such a complex matter that it is difficult to assess by disaggregating it. Synthesis may have to be assessed holistically through Assessment by Professional Judgment (Vars 2002). Those who seek to assess synthesis analytically may need to look for the footprints of synthesis rather than for synthesis itself. There are many possible footprints of synthesis; these are often referred to as *indicators* in assessment literature.

How May We Collect Data on the Achievement of Interdisciplinary Learning Objectives?

The AIS Assessment Committee recommends collecting data representative of a variety of perspectives, and that an emphasis on performance data be included in any plan. It is beyond the scope of this article and perhaps too prescriptive to give specific recommendations for assessment techniques, but a general outline of strategies follows. The most important criterion in choosing the types of data to collect is whether or not the data will provide useful information on what students are learning.

Indirect Assessment Techniques of Interdisciplinary Assessment

Indirect assessment data are created when students reflect on their own learning rather than actually exhibit direct evidence of an educational outcome. As defined earlier, indirect data may be obtained from various environments, such as program questionnaires, exit interviews, focus groups, and alumni surveys (Palomba & Banta 1999, pp. 11-12). The survey in one incarnation or another is the most frequently mentioned platform for gathering indirect data on interdisciplinary learning.

Direct Assessment of Interdisciplinary Learning

There are numerous commercially available instruments on the market today that may aid in assessing a particular interdisciplinary program. While some of these instruments are highly developed surveys, and thus fall in the domain of indirect assessment, others take the form of tests that yield direct, proxy evidence of learning. Standardized and locally developed tests both have potential to inform interdisciplinary teaching and learning. We find that

most commercially developed instruments do speak to some of the intended outcomes of interdisciplinarity—critical thinking and argument analysis, for example—but there is a real danger of defaulting program outcomes only to those measured by the instrument.

Performance Assessment

Evolving assessment wisdom favors an emphasis on direct data—data created as the actual products and behaviors of students. Performance data for direct assessment are created when students display their knowledge, skills, and attitudes by responding directly to prompts in such environments as complex assignments, internships, projects, service learning, and real and simulated problems. These types of assessments are often referred to as authentic or performance assessments. In interdisciplinary programs that seek to “develop complex integrated skills,” faculty members often find performance assessment appealing because it asks students to display their skills in a way that is more thorough and nuanced than that provided by traditional paper and pencil tests (Palomba & Banta 1999, p. 116). Performance assessment is proving to be an extraordinarily rich environment upon which a number of data-gathering perspectives may be imposed. Wiggins (1998) describes performance assessment as gathering data in the kind of work real people do rather than provoking easy to score responses to simple questions (p. 21). Wiggins’ criteria for a performance assessment require that the tasks:

- be realistic;
- require judgment and innovation;
- require the student to perform the subject;
- replicate real-life contexts;
- assess student ability to efficiently and effectively use a range of knowledge and skills to negotiate a complex, real-life task; and
- allow some opportunity to rehearse, practice, and receive feedback on one’s performance. (pp. 23-24)

One type of performance assessment is the senior assignment, which is a scholarly engagement between student and dedicated faculty member(s) that results in a product or behavior (Eder 2001, p. 201). Because the product or behavior is visible, it, and the curriculum that produced it, can be assessed. Thus, the senior assignment provides an environment that makes student achievement visible as a valid indicator of student learning. Over the course

of time and experience, the senior assignment is an excellent vehicle for linking an authentic learning environment to a program's curriculum and pedagogy (Wiggins 1993a).

The development and application of rubrics in performance environments is encouraging, and we recommend their application to interdisciplinary programs. A rubric is a visible standard—and a scoring guide—that allows the assessor and the public, for that matter, to recognize expectations and to make increasingly fine distinctions about the quantity and quality of student learning. A particularly powerful assessment device that invokes rubrics is Primary Trait Analysis, or PTA (Walvoord & Anderson 1998). Whereas faculty members often balk at assessment, PTA takes something they already do, namely grading, and turns it into assessment. A particular advantage of PTA is that it operates well in many environments, from single assignments and classes to entire programs and curricula. Many recent publications and workshops at professional organizations focus on PTA and the nuances of rubric development (Montgomery 2001). Suffice it for now to suggest that in regard to interdisciplinary programs, rubrics:

- are particularly useful in assessing complex intellectual processes;
- promote objectivity, reliability, and validity in assessment;
- clearly specify what the student is expected to learn;
- clearly specify what is to be taught;
- provide opportunities for assessment in the course of the semester, thus allowing adjustments to be made; and
- provide useful information on teaching and learning.

When May We Collect Data on Our Learning Outcomes?

One of our biases is that IDS assessment is primarily formative and, because of the diversity of IDS programs, IDS assessment is local. It is up to the will of the faculty to determine when and where information will be collected. Ideally, assessment will be conducted regularly and appropriately. The complexity of the assessment can suit its time and purpose. Assessment in a course may be as simple as the instructor conducting a simple Minute Paper (Angelo & Cross 1993) or a focus group after the first two weeks of class to determine what and how well the students are learning. A more complicated, yet equally useful, assessment may occur at the end of the term as the students engage in some authentic performance of their learning. The information obtained at these points can assist in grading as well as in the preparation for teaching

the class the next term or even the next day. Assessment information of all types should be reviewed continuously or, more practically, periodically, to insure that learning is occurring and to make appropriate adjustments when such is not the case.

What Will We Do with the Information We Collect?

Information without meaning has no power. The central conceit of many assessment programs is that they have entered the assessment arena motivated by external necessity. As accreditation or self-study looms large, assessment becomes a hot topic. The mistake made by many is that they will develop a promising assessment plan, gather some data, present those data in fine form, mention some changes anticipated as a consequence of the data, and receive re-accreditation. Breathing a sigh of relief, they announce that they do not have to worry about assessment for another ten years, or at least until another regulatory report is due. The single biggest acknowledged failure of assessment, both in terms of its performance and being taken seriously by skeptics, is the absent application of relevant, timely feedback from otherwise noble expenditures of time, energy, and resources. What a tragedy that such a noble reform effort becomes established as such an onerous task.

Assessment requires a fundamental change in the way we think about interdisciplinary education and go about our daily tasks. A pilot flying from New York to Los Angeles must constantly seek data, make decisions, and refine flight plans as a consequence of a dynamic environment. The teaching-learning process is no less dynamic, yet we are often inclined to treat it as the most static of endeavors.

Fundamental Principles of Interdisciplinary Assessment

Before we conclude, it will be useful to take some general assessment principles and situate them in an interdisciplinary environment. We have developed some reasoned positions that we hope will encourage interdisciplinarians to move ahead with assessment. Hopefully these thoughts will become the subject of much reasoned discourse as the interdisciplinary community incorporates assessment into the existing models. The following principles, though described in the interdisciplinary context, correlate with the works of others in assessment (Wiggins 1998, pp. 17-18).

- Assessment of interdisciplinary programs should be an ongoing, systematic process that produces real results for improve-

ment of teaching and learning. Assessment is not a consuming affliction from without, but rather a tool to be used judiciously by faculty members. It need not continuously consume our resources or our attention. Indeed, assessment works best when it is applied at key points and times in the curriculum. As such, assessment is a central component of effective interdisciplinary instruction. (Field & Stowe 2002)

- Assessment of interdisciplinary programs should arise from a union of purpose, not from the available techniques and tools of assessment. A major purpose of interdisciplinary studies is to help students gain integrative knowledge, skills, and perspectives. A major purpose of assessment is to help the student to learn better and the teacher to instruct better. Assuredly, these two purposes converge regardless of which pedagogical techniques and assessment tools are chosen.
- External accountability activities, such as performance funding and benchmarking, must not be allowed to confuse or supersede the primary internal purpose of assessment.
- Interdisciplinary assessment is most effective when conceived locally, in concert with a well-defined mission, employing a variety of measurement perspectives and techniques, both direct and indirect. To be sure, external guidelines from funding and accrediting agencies will inform desired outcomes, but it will always be a local cycle of performance and feedback that enhances learning and teaching. Interdisciplinary mission is local; therefore, interdisciplinary assessment is local.
- Just as assessment is a central component of instruction, performance tasks are a central component of a comprehensive interdisciplinary assessment program. (Eder 2001)
- Assessment and evaluation are related, but clearly different notions. Even though their processes are similar, one must view assessment and evaluation, not merely from the perspectives of their processes, but also from the perspectives of their outcomes and consequences. The major consequence of assessment is continuous improvement of teaching and learning. The consequence of evaluation is the ability to make an informed decision about worth or value. Mandated external evaluation should not be confused with, or replace, assessment.

Conclusion

No doubt one's perspective on assessment, particularly the assessment of interdisciplinary programs, will change through actual assessment experience. With some experience we know that an overly rigid reliance on testing and measurement in interdisciplinary assessment has not fulfilled the promise of assessment. A final definition of assessment offered by Wiggins (1993b) seems more appropriate now: "assessment is a comprehensive, multifaceted analysis of performance; it must be judgment based and personal" (p. 13). Measurement is important, but assessment gives meaning to measurement. Assessment should begin the discourse about interdisciplinary teaching and learning; it should not lead to the end of the story.

Assessment of interdisciplinary programs is a process that requires the active participation and goodwill of stakeholders who are mutually committed to improving teaching and learning in a variety of interdisciplinary environments. Assessment of interdisciplinary programs is fundamentally a people process. It is secondarily a technical process, and the two should not be confused. Primacy should be given to the involvement of multiple stakeholders. IDS programs that "view assessment as just a process or as just a set of technical tasks will either risk basing their decisions on faulty information or will have difficulty implementing any results" (Erwin 1991, p. 16).

Biographical note: Don Stowe is the Associate Dean of the College of Hospitality, Retail, and Sport Management at the University of South Carolina. He currently serves as the Chair of the AIS Assessment Committee.

Douglas Eder is an Associate Professor of Neuroscience and Director of Undergraduate Assessment and Program Review at Southern Illinois University—Edwardsville. He is a member of the AIS Assessment Committee and a national leader in the assessment movement.

References

- Angelo, T.A. (1995a, April). Reassessing assessment: Embracing contraries, bridging gaps, and resetting the agenda. *AAHE Bulletin*, 47(8), 10-13.
- . (1995b, November). Reassessing (and defining) assessment. *AAHE Bulletin*, 48(2), 7-9.
- Association of American Colleges and Universities. (2002, October). Greater expectations: A new vision for learning as a nation goes to college. Washington, DC: Author.
- Astin, A.W. (1991). *Assessment for excellence: The philosophy and practice of assessment and evaluation in higher education*. New York: American Council on Education.

cil on Education/Macmillan.

- Baxter-Magolda, M.B. (2001). A constructivist revision of the measure of epistemological reflection. *Journal of College Student Development*, 42(6), 520-534.
- Biggs, J. (2002). Assessment: Where did we go wrong? A reply. *Assessment Update: Progress, Trends, and Practices in Higher Education*, 14(4), 6-7, 15.
- Dickeson, R.C. (1999). *Prioritizing academic programs and services: Reallocating resources to achieve strategic balance*. San Francisco: Jossey-Bass.
- Eder, D.J. (2001). Accredited programs and authentic assessment. In C.A. Palomba & T.W. Banta (Eds.), *Assessing student competence in accredited disciplines* (pp. 199-216). Sterling, VA: Stylus Publishing, LLC.
- Eisner, E. (1998). *The enlightened eye: Qualitative inquiry and the enhancement of educational practice*. Upper Saddle River, NJ: Merrill/Prentice Hall.
- Erwin, T.D. (1991). *Assessing student learning and development*. San Francisco: Jossey-Bass.
- Farmer, D.W., & Napieralski, E.A. (1997). Assessing learning in programs. In J.G. Gaff, J.L. Ratcliff, & Associates (Eds.), *Handbook of the undergraduate curriculum: A comprehensive guide to purposes, structures, practices, and change* (pp. 591-607). San Francisco: Jossey-Bass.
- Field, M., & Stowe, D. (2002). Transforming interdisciplinary teaching and learning through assessment. In C. Haynes (Ed.), *Innovations in interdisciplinary teaching* (pp. 256-274). Westport, CT: American Council on Education/Oryx Press.
- Klein, J.T. (2002). Assessing interdisciplinary learning K-16. In J.T. Klein (Ed.), *Interdisciplinary education in K-12 and college: A foundation for K-16 dialogue* (pp. 179-196). New York: The College Board.
- Montgomery, K. (2001). Authentic tasks and rubrics: Going beyond traditional assessments in college teaching. *College Teaching*, 50(1), 34-39.
- National Center for Postsecondary Improvement. (1999). Revolution or evolution? Gauging the impact of institutional student assessment strategies. *Change*, 31(5), 53-57.
- National Research Council. (2001). *Knowing what students know*. Washington, DC: National Academy Press.
- Newell, W.H. (1998). Professionalizing interdisciplinarity: Literature review and research agenda. In W.H. Newell (Ed.), *Interdisciplinarity: Essays from the literature*. New York: The College Board.
- Palomba, C.A., & Banta, T.W. (1999). *Assessment essentials: Planning, implementing, and improving assessment in higher education*. San Francisco: Jossey-Bass.
- Schilling K.M., & Schilling, K.L. (1998). *Proclaiming and sustaining excellence: Assessment as a faculty role* (ASHE-ERIC Higher Education Rep. 26[3]). Washington, DC: George Washington University Graduate School of Education and Human Development.

- Schneps, M.H., & Sadler, P.M. (Director), & Harvard Smithsonian Center for Astrophysics (Producer). (1987). *A Private Universe* [Video Documentary]. (Available from Annenberg/CPBP Pyramid Film and Video, <http://www.learner.org/catalog/series28.html>).
- Scriven, M.S. (1974). Pros and cons about goal-free evaluation. *The Journal of Evaluation Comment*, 3(4), 1-4.
- Squires, G. (1992). Interdisciplinarity in higher education in the United Kingdom. *European Journal of Education*, 27(3), 201-210.
- Tyler, R.W. (1949). *Basic principles of curriculum and instruction*. Chicago: University of Chicago Press.
- Vars, G. (2002). Educational connoisseurship, criticism, and the assessment of integrative studies. *Issues in Integrative Studies* 20, 65-76.
- Walvoord, B.E., & Anderson, V. (1998). *Effective grading: A tool for learning and assessment*. San Francisco: Jossey-Bass.
- Wiggins, G.P. (1993a, November). Assessment: Authenticity, context, and validity. *Phi Delta Kappan*, pp. 200-214.
- . (1993b). *Assessing student performance: Exploring the purpose and limits of testing*. San Francisco: Jossey-Bass.
- . (1998). *Educative assessment: Designing assessments to inform and improve student performance*. San Francisco: Jossey-Bass.