



UNIVERSITY AT ALBANY
STATE UNIVERSITY OF NEW YORK

**ASSESSING STUDENT ATTAINMENT
IN THE ACADEMIC MAJOR**

WHAT'S THE QUESTION ?

Research Report No. 16

Dr. J. Fredericks Volkwein
Director, Institutional Research and Faculty Member in
Educational Administration and Policy Studies

University at Albany
State University of New York

ASSESSING STUDENT ATTAINMENT IN THE ACADEMIC MAJOR

This paper outlines the philosophical foundations for assessment of student learning, summarizes the generic purposes of collecting evaluation and assessment information, and offers a list of methodologies for academic departments to consider.

Philosophical Foundations

There are two Philosophical foundations for our approach to assessment at Albany: One inspirational, the other pragmatic.

1. The inspirational: Simply put, our goal is the enhancement of student learning and growth. For two decades, Albany has carried out assessment and self-evaluation not to please outsiders, but to satisfy ourselves. We have undertaken assessment not to judge undergraduate education, but to improve it. Thus, assessment is not a product or an end, it is a process or a beginning.

Our goal has been to achieve a formative, rather than a summative faculty attitude -- an organizational climate of cooperative development and on-going improvement, rather than authoritative finality.

2. The Pragmatic: In an atmosphere of scarcity, those academic departments able to provide evidence about the impacts they are having on their students, will most assuredly do better in the competition for resources than academic departments that are not able to provide such evidence.

These dual emphases, these Janusian purposes, seem to offer a constructive path. They provide information for our internal development, at the same time giving us a better means of competing for scarce resources.

FOUR ASSESSMENT QUESTIONS

Evaluation and assessment force us as professionals to engage in evidence-based thinking. Moreover, the nature of the evidence we gather depends upon the question one asks at the beginning of the process. Most evaluation and assessment activities seek answers to one or more of these generic questions: Do we meet the standard? How do we compare? Are we meeting our goals? Are we getting better?

Does the student or program meet or exceed certain standards?

The answer to this question requires what my measurement colleagues refer to as "Criterion Referencing Assessment." This is a form of Summative Evaluation that is most apparent in the traditional accreditation and certification approaches where students and programs are judged by the whether or not they meet certain standards that have been agreed upon by faculty or other experts in the field. This assessment approach requires definition and agreement about minimum standards or threshold levels of performance. While true scientific measurement is rare, tests of competency and performance skill, quantitative measures of knowledge, licensing exams, and professional judgment methodology are common. While the performance standards for individual students can be locally determined (such as a basic skills test), this type of assessment usually involves the employment of state (the Bar Exam), regional (Middle States), or national (Medical Boards) standards. Centralized approaches, therefore, are the most common.

Advantages

- Focuses on ensuring minimum competencies among all students and programs.
- Requires agreement about acceptable levels of performance, thus clarifying the educational and curricular objectives.
- Useful for making summative decisions about the continuance and discontinuance of students and programs.

Disadvantages

- Largely ignores how much students have learned, and few faculty are comfortable disregarding where students started from in the learning process.
- Agreement about acceptable levels of performance may be difficult to achieve.
- Little basis of comparison with performance at other institutions, except on the basis of the percent of students that pass.
- Like pass/fail grading, such assessment often make no distinction between acceptable and top performance.
- Basic validity rests with the assessment instrument and the definition of the performance criterion (this can be a problem especially with locally designed instruments).
- If the standard is set externally, the curriculum loses its local focus.

How does the student or program compare to others?

Another form of summative evaluation, this approach is referred to as "Norm Referencing Assessment" in the research literature. Performance is judged in relation to that of an appropriate comparable group of students, departments, or institutions. To address this question requires comparative and relative measurement, usually using national or state instruments and norms. A fairly centralized approach is normally required. Examples might include the E.T.S. Major Field Achievement Tests, G.R.E. scores, and reputational surveys. The following advantages and disadvantages characterize this approach.

Advantages

- Gives opportunities for comparison with students and programs at similar institutions.
- Using national and state comparisons has more external legitimacy.
- Requires no consensus about performance standards.
- Like "grading on the curve," this is perhaps a useful approach if one is forced to select or screen a reduced number of students due to limited resources or excess demand.

Disadvantages

- Requires the identification of appropriate normative reference groups and raises the potential of their non-comparability.
- Nationally designed tests may or may not be congruent with local curricular objectives.
- Gives less information about trends and about levels of attainment, especially if the scores are relational (this is a serious deficiency of the GRE).

Does the student or program do a good job at what it sets out to do?

Measuring goal attainment, or "Internal Referencing Assessment," requires clear goals and educational objectives. Unlike the first two approaches, this question requires a more formative student-centered emphasis. This approach often attempts to measure the congruence between objectives and actual performance and uses quantitative evidence only to the extent that goals and objectives can be quantified. More commonly, professional judgment evidence is utilized.

Advantages

- Requires clear goals and educational objectives.
- Recognizes the longitudinal and formative nature of the educational process.
- Is easily tailored to the local curriculum and the needs of particular students.
- This is a relevant question for most students and institutions.
- Likely to be an effective form of evaluation.

Disadvantages

- Likely to be an inefficient form of evaluation requiring extra faculty effort.
- Provides little useful summative information to external constituencies.

Has the student or program improved and How can the student's program and learning experience be improved even further ?

This approach is the most formative and developmental of all and is most noticeably practiced at places like Alverno College and Kings College. This is the one question that is universally applicable to ALL students, institutions, programs, and faculty. Multi-method, multi-measure, and decentralized research strategies are most fruitful, and departments should be encouraged to use a variety of assessment tools and sources of evidence. The following advantages and disadvantages characterize this approach.

Advantages

- Applicable to all students and programs regardless of potential and previous attainment.
- Incorporates a wide and flexible range of objective and subjective sources of information.
- Maintains a focus on the student and the learning process.

Disadvantages

- Requires continuing faculty attention and periodic renewal.
- Improvement can be difficult to demonstrate in the absence of a pre-test/post-test methodology.

Alternative Methods for Assessment in the major

At Albany each department constructs an appropriate means for assessing student attainment in the major using a broad range of possible assessment designs. In some departments where the numbers of majors are small, the faculty might focus on the achievement of all students. Other departments, perhaps with large numbers of majors, may decide to study representative groups of students. The aim of the assessment is not to magnify any particular student's success or failure, but rather to judge in whatever ways we can the students' abilities to construct knowledge for themselves and to assess our work as faculty in nurturing, enhancing, and enabling those abilities. Departments might combine some of the elements discussed below, or might design a unique approach, not considered here. Our objectives are to examine what it is students are acquiring when they major in a particular discipline, and to use that information to enhance the learning experience for future students.

ALTERNATIVES:

Comprehensive Exam

When such exams are designed locally by the faculty they have the advantage of being shaped to fit the department's curriculum. Departmental exams have the disadvantages of needing labor-intensive annual revision and local scoring by the faculty, and of lacking a comparison group. Standardized instruments provide scoring services and more reliable and valid comparison groups; but they may or may not fit the department's curriculum, and are not useful in disciplines wishing to go beyond a multiple choice format.

Senior Thesis or Research Project

Such a requirement encourages students to use the tools of the discipline on a focused task as the culmination of their undergraduate academic experience. Under ideal conditions, each department or program uses the student's work to reflect on what students are achieving with the aim of evaluating, and if necessary strengthening, the curriculum and experiences of students within that major.

Performance Experience

Asking students to demonstrate in some practical, or even public way, the knowledge and skills they have learned and acquired, this emphasizes the integration and application of the separate facets of the academic major. Such a requirement may be especially fitting in professional and performing arts fields. Examples include student recitals, art exhibitions, practice teaching, and supervised field experiences.

Capstone Course

This is usually a required senior course designed to integrate the study of the discipline. It often has a heavy research and writing component. Such courses offer ideal opportunities, both to assess student learning and to strengthen the curriculum of that major. Often the course can contain some other form of assessment, such as a local or national exam, embedded within it.

Student Portfolio of learning experiences

In this mode of assessment, students collect systematically the work that they have undertaken in their study of a discipline. They undertake and write a self-examination of the material, demonstrating how they have constructed the discipline through their writing and thinking over two years of study. Afterwards, faculty meet with students to go over this portfolio. The faculty then could use their own and the students' analyses of portfolios coupled with their perceptions of the student conferences as a basis for conversation among faculty about the curriculum and practices within a discipline. In departments selecting this option all faculty responsible for undergraduate education should be a part of this process, but the plan becomes difficult to implement if each faculty member has to assess overly large numbers of students.

Senior Essay and Interview

The faculty constructs a series of questions that ask students to demonstrate their conceptual understanding of the discipline, and to reflect on the strengths and weaknesses of their programs. The students respond in writing and then meet with faculty to discuss their written statements. The faculty then meet to discuss the results of their conferences with students for the purpose of strengthening the major. Large departments may sample a cross section of seniors rather than the entire population.

Other suggestions:

- **Course embedded assessment** (e.g., standardized examination; classroom research project; beginning vs. ending written assignment)
- **Student Self-Assessment** (e.g., video of oral presentations)
- **Student Peer assessment** (builds a learning culture in the department)
- **Secondary reading** of student's regular course work by faculty colleagues (requires an atmosphere of department collegiality)
- **External examiners/readers** (puts student and faculty member on the same learning team - they both want to look good.)
- Using **historical department data** on student performance and placement
- Faculty **Focus Group** discussions with students
- **Surveys of students and alumni using self-reported measures**

SUMMARY and CONCLUSIONS

Table 1 gives a summary of possible methods for assessing student attainment in the major depending upon each purpose. The suggestions contained in the table should be treated as guides for departmental consideration, rather than as definitive uses. A capstone course or departmental exam might be the most useful for assessing the student's attainment of faculty determined performance standards. For the purposes of meeting professional or national standards, a competency test or external examiner may be employed. Comparing a student's performance to some external reference group almost always requires the use of a national exam or performance standard, but even graduate school and placement information can be used for comparative purposes. Measuring the attainment of a student's educational goal requires student-centered strategies such as learning portfolios or focus groups. To answer the question about getting better, a wide range of strategies using multiple sources of information can be utilized.

By now, it may be obvious to the reader that each assessment question or purpose can be addressed in a variety of appropriate, yet different ways. The question of whether the activity is relatively centralized and controlled by forces outside the department, or decentralized and controlled by the faculty, is perhaps less important than the usefulness of the assessment for enhancing student learning. Even standardized, nationally normed tests can be used by faculty to reflect on the strengths and weaknesses of the department's program. And even the most student-centered form of talent development usually yields results that can be aggregated to serve program evaluation purposes.

Judgments of student learning and program effectiveness will be better and richer if they are based upon multiple indicators and measures, than if based upon a single indicator or measure. First, the use of a single measure (such as a single test) is likely to be unfair to students either who prepared heavily in an area not covered by an exam or who had an "off" day. Second, a multi-dimensional approach yields a far more reliable image of strengths and weaknesses. For example, when a department sees congruence between the information it receives from graduating seniors' test results and from alumni survey results, it has more confidence in the strength of the findings.

There are many appropriate uses and efficiencies to be gained from relatively centralized data collection activities carried out in campus offices of institutional research. Institutional researchers often work with others on the campus to develop and maintain student and alumni information systems that can serve as rich assessment databases. Rather than developing their own survey research, database, and software expertise, departments can draw upon these centralized databases in conducting their own assessment activities.

Several Albany Departments serve as good examples of cooperative data sharing between administrative offices and academic departments. We are able to download from our centrally-maintained student and alumni databases a variety of useful information that is helpful to student and program assessment by faculty.

The Albany alumni studies is an excellent example. The cost to the institution of conducting separate alumni surveys in each department and maintaining independent databases is too high. Therefore, during the 1980s we met with several of our largest departments and designed a survey containing items of value for both centralized and decentralized purposes. After surveying several thousand alumni, we are now in a position to distribute both aggregated and disaggregated responses to a wide range of measures of alumni satisfaction, educational outcomes, intellectual growth, career development, and effectiveness of the Albany undergraduate experience, both departmental and campus-wide.

An important ingredient in a successful assessment program is an attitude of cooperation and trust among faculty and between faculty and administrative staff. Faculty need to be trusted to use the information for the enhancement of student learning, and administrators need to be trusted to use the information to promote institutional effectiveness. Assessment evidence should not be used for faculty evaluation. Assessment promotes evidenced-based thinking and new conversations.

TABLE 1
ASSESSMENT IN THE UNDERGRADUATE MAJOR FIELD OF STUDY

<u>Four Purposes</u>	<u>Suggested Strategies and Methods</u>
Meeting Standards	<ul style="list-style-type: none"> Department Comprehensive Exam Capstone Course Senior Thesis/Research Project External Examiners/Readers Student Performance/Exhibit Practice Teaching Competency Tests Certification Exams
Comparing to Others	<ul style="list-style-type: none"> National Comprehensive Exam Peer Assessment Program Review by External Experts Comparative Graduate School and Placement Informat.
Measuring Goal Attainment	<ul style="list-style-type: none"> Senior Thesis/Research Project Student Portfolio Student Self-Assessment Focus Groups Alumni Surveys Exhibits and Performances
Improving and Developing Talent	<ul style="list-style-type: none"> All of the Above, especially: Senior Essay and Interview Focus Groups Course Embedded Assessment Learning Portfolio Alumni Surveys External Program Reviews