

# Report of the ADE Ad Hoc Committee on Assessment

THE following report was approved by the ADE Executive Committee at its March 1996 meeting. Members of the ad hoc committee: David Bartholomae, University of Pittsburgh, Pittsburgh; Terrell Dixon, University of Houston, University Park; Pat Hutchings, American Association for Higher Education; David Laurence, ADE; Carol Martin, Boise State University; Stephen Olsen, ADE; Alan Purves, State University of New York, Albany; and Judith Wilt, Boston College (chair).

## Charge to the Committee

The committee was charged to investigate the issues and procedures of assessment in English and to consider what advice the ADE can usefully offer to departments and chairs engaged with the problem of developing assessment initiatives, especially initiatives focused on the documentation of student learning. In carrying out this charge, the committee sought information from chairs of ADE member departments, consulted with colleagues at ADE Summer Seminars and the MLA Annual Convention, and formulated advice regarding departmental practice in this area.

## Introduction

Over the years the Association of Departments of English has made recommendations for good practice in a number of areas suggested for timely study by members. Assessment is one of the most complex and controversial of these areas. When the ADE ad hoc committee asked chairs of ADE member departments to comment on their experience with assessment, we heard vivid stories of hope, challenge, and frustration—of academic life sometimes demoralized where it was meant to be renewed, sometimes invigorated where it was feared only boilerplate would be produced. We heard too from those who thought nothing need be said yet at all about this still tender and conflicted topic. And we were very much alert to the prospect that assessment may in some circumstances be used for purely bureaucratic or downsizing agendas void of real academic purpose or promise.

But as a profession we are unmistakably called to this vexed intersection of community oversight, professional development, and academic freedom. The ad hoc committee is persuaded that faculties of ADE's member departments will do best to arrive at that intersection with at least some of the fruits of shared experience and reflection. In this context, the early and ongoing experience with assessment of writing has been a valuable starting point both for the committee and, clearly, for many of the departments now engaged in assessment activities.

Almost by definition, assessment is destined to meet a skeptical reception from faculty members in English, at least initially. Departments commonly formulate assessment plans because they must meet requirements imposed by agencies beyond their campuses—most often the regional accreditation commission for the institution, on occasion a state board of regents, education department, or legislature. Often the motives behind assessment are perceived to be accountability and efficiency—the bureaucracy's search for easily grasped measures of program performance on which to base decisions about the allocation of resources. One of the founding premises of the assessment movement has been that the quality of an educational program rests primarily on what students learn as a result of the instruction given by that program. In bureaucratic contexts "what students learn" can easily metamorphose into a simple quantitative increment of "value added." For a few of its champions and many of its detractors, assessment amounts to little more than a search for instruments, processes, or methods that will capture "program quality" in a single number indicative of what a given population of students learned as a result of undertaking a certain course of formal study.

However, it is equally important to recognize that the assessment movement is broader, more thoughtful, and of considerably greater interest than the caricature to which proponents and detractors alike occasionally reduce it. Assessment's concerns are far from new. Indeed, in asking about the cumulative effects of college and about a given student's readiness for a degree, assessment in some ways recalls earlier solutions to the problem of educational quality. In the nineteenth-century American

college it was assumed that the completion of courses was by itself insufficient warrant for a degree. Candidates for the bachelor's, therefore, faced a final hurdle in the senior declamation, a requirement that they publicly demonstrate possession of knowledge and high skills of intellect and speech. Earlier in this century comprehensive examining of seniors played a role in the awarding of degrees. Such practices exemplify a tradition of asking questions about graduates' competence and about the cumulative effects of teaching and curricula.

In important ways the impetus for assessment arises inside the profession; it is not solely imposed from outside. Evaluation, the close and continual appraisal of individual students in particular courses, is central to the teaching enterprise. As the study in common of the students a department's faculty teaches in common, assessment extends the tradition of reflective self-study carried on by groups within the university community even as it guides the scrutiny of our professional activity by the wider community the profession serves. Viewed from this angle, it holds promise as an exercise of collegiality directed toward the improvement of our educational activity.

Within the academy assessment has been a distinct field of scholarly inquiry for more than two decades. Teacher practitioners in secondary and higher education have profited from its insights, its instruments, and especially its cautions about its insights and instruments. During the 1980s a new force entered the field, as governors and legislators, corporate members of boards of trustees, and cost-conscious citizens claimed the right to be heard, or even to prevail, in the common shaping of the future of secondary and higher education. As Pat Hutchings and Ted Marchese note, in 1986 the governor of Missouri, John Ashcroft, as chair of the National Governors' Association Task Force on College Quality, bluntly stated the links among assessment, accountability, and resources:

The public has a right to know and understand the quality of undergraduate education that young people receive from publicly funded colleges. . . . They have a right to know that their resources are being wisely invested and committed. . . . We need not just more money for education but more education for the money. (qtd. in Hutchings and Marchese 16)

Since 1986 some forty states have enacted assessment mandates of one sort or another for their institutions of public higher education. Some of these mandates are legislative; most come from governing or coordinating boards in state education departments; a few arise at the level of state public higher education systems. In some states and systems these mandates have recently been pushed into the background by state budgetary problems. But the questions that prompted the mandates—about

how higher education can document that it is working responsibly and effectively, especially in its programs of undergraduate education—have, if anything, intensified.

These questions also stand behind the call for assessment that has come from regional accrediting agencies. The United States secretary of education has the authority formally to recognize the various accrediting associations that together make up higher education's voluntary accreditation system; the secretary also has the power to issue further regulations governing the associations' activities. (For a list of the regional and national accrediting agencies formally recognized by the secretary of education, see the *1996 Higher Education Directory* vi–xiii.) Since 1988 these regulations have included a requirement that both regional and specialized accrediting bodies require the institutions they accredit to supply “evidence of academic achievement”—that is, that the institutions conduct assessment. Unlike state mandates, which typically apply to public institutions only, accreditation affects virtually all institutions, private and public alike, particularly since under the Higher Education Act institutions must be accredited to participate in federal student aid programs. For English departments, the seven regional accrediting associations are in fact the salient source of pressure for assessment on most campuses.

But meanwhile—contributing to the ambiguities of assessment—the status and fortunes of the voluntary accreditation system have come into question. For several years, university and college presidents have expressed their increasing restiveness with the multiplying demands imposed by the various accrediting bodies to which colleges and programs in their institutions are subject, and reform efforts under way are aimed at bringing to the accreditation system greater coherence, coordination, and credibility. On another front, the unsettling course of congressional debate during the reauthorization of the Higher Education Act in 1992 raised concerns that direct state and federal regulation of higher education could follow should the voluntary accreditation system prove incapable of strengthening itself. Indeed, in 1993 federal legislation calling for the creation of State Postsecondary Review Entities, which would have broad investigatory powers and would be administered by the states, was on the verge of being implemented. (For further information about the developments briefly alluded to here, see Atwell; Longanecker; Warren; and Benjamin. See also Poston.) As of this writing, in the spring of 1996, the practical consequences of the movement to reform the voluntary accreditation system remain uncertain. Whatever proposals finally emerge, the committee is persuaded that departments can act to their own best advantage by taking steps that put them in a position to document what the teaching they provide does for students. In the committee's view, assessment conducted in the context of

voluntary accreditation represents an alternative far preferable to direct regulation by federal or state agencies.

The committee does not minimize the tensions that exist between the perspectives policy makers bring to the question of assessment and the views of faculty members in our community. Traditionally, assessment in English courses most frequently involves some sort of writing assignment, and members of our community know that the quality of writing people produce is closely linked to the specific means through which, and the specific occasions and purposes for which, performances are solicited. Further, faculty members are keenly aware that these context-specific contingencies are subject to considerably greater variation than a distant observer might assume. The intellectual values informing the way faculty members work with students emphasize the individual, idiosyncratic aspects of literary interpretation and writing performance, as well as the conundrums and complexities that arise in connection with evaluation in literary study. In many respects the approaches to evaluation commonly used in our community make a poor fit with the values governing formal assessment across multiple courses or course sections, departments, or institutions, particularly the values embedded in technical standards of validity, reliability, and comparability. The information contained in the performance-based assessments faculty members tend to find most interesting and helpful—portfolios, interviews, locally prepared exams, senior capstones, surveys, diagnostic writing samples, and the like—almost certainly needs careful reformulation if the public policy arena is to recognize it.

In the light of these tensions, special attention must be given to the kinds of evidence and manner of presentation departments expect to rely on if they are called on to create assessment plans. What can stand as credible evidence of student learning? What forms of evidence are readily available in departments already? And what languages are available or being developed—in literary criticism or rhetoric or cognitive psychology or elsewhere—to articulate these evidences and communicate them to relevant audiences in a persuasive and credible manner? The committee's sense is that faculties in English, since their members are trained in occupying multiple standpoints and perspectives, are often in an advantageous position to engage these questions fruitfully.

#### **Assessment in English: Comments from the Field**

In February 1994 the ADE ad hoc committee canvassed ADE member English departments about the range of departmental assessment practice and experience with assessment. The committee asked chairs to comment particularly about their experience with undergraduate-

program assessment that calls for the documentation of student learning. Specifically, the committee asked the following questions:

- Has your department been required, or will it soon be required, to develop some assessment of its undergraduate major and/or general education programs? As part of the assessment are you required to document what students, considered as a group, experience in and gain from the program?
- From what source or sources is the assessment requirement coming: a state mandate, the policies of a regional accrediting body, or some other source (e.g., your own institution's or department's self-study, or questions from parents of prospective students or the board of trustees)?
- What specifically have you been asked to do or provide? Have faculties of individual departments in your institution been asked to develop an assessment strategy as they see fit? Have you been provided any written guidelines? Has a distinction been made between the undergraduate English major and general education programs in writing and literature? Has any specific test or kind of test been required?
- Do you have a sense of the kinds of evidence you are being asked to provide and of the kinds of evidence that are regarded as acceptable or credible? What is the source of your understanding?
- Have activities connected with assessment of student learning had any effects on faculty members' attitudes and behavior or on departmental policies and curriculum?

Chairs of ninety-five departments (just over ten percent of the membership) responded, many in generous and helpful detail. While such a self-selected sample cannot provide a basis for systematic analysis and generalization about attitudes and practices nationally, the chairs' letters do serve as a useful guide to the range of English departments' experiences with assessment and of approaches to it. Twenty-five chairs indicated that assessment had been mandated and was already under way on their campuses, and another thirty-seven reported that assessment had just begun or was about to begin. Chairs of nine departments said assessment was possible or expected on their campuses in the near future. Chairs of twenty-four departments reported that assessment had yet to reach them or their campuses.

Unsurprisingly, chairs of departments where assessment mandates were, or soon were to be, in effect reported ambivalence about the requirements. Persuading faculty members—and themselves—to take assessment seriously and approach positively the tasks involved in it was a primary challenge for many. Yet many chairs also noted that assessment, if appropriately conceived and implemented, can intersect with faculty members' long-standing concern for the quality of their teaching and programs and with the field's traditions of departmental self-study. As one chair wrote, he and his colleagues hoped "that the practice of assessment will make us all more self-conscious about

our classroom practices, that it will involve us in healthy, continuous self-scrutiny, and that it will help us keep our programs coherent and purposeful.” Another chair urged that “we ought to regard [assessment] not as a silly administrative bookkeeping burden but as an opportunity to demonstrate our successes.”

Most of our correspondents reported anxiety and resistance among department members, especially in the beginning. Of special concern for our correspondents was the haste with which assessment mandates were in some instances supposed to be implemented. One chair vividly described the questionable results obtained when an institution hurries pell-mell to create an assessment plan without making adequate provision for thought and professional development.

We had only one “educating” event (a speaker) to fill us in on assessment, what it means, how to do it. He seemed to stress taking time with the process, which we have not done. . . . I wish we had had time to talk with chairs at other small liberal arts colleges to find out what measures they are using and to read more research on what are valid measures in our field. ADE guidelines on this would have been very useful.

Faced with a sudden imperative for departmental assessment, this chair turned to a branch of the discipline that has had early and continued experience with assessment: “I concocted the portfolio system because of what I’ve heard and read about it via the state Writing Project.” Though departmental assessment programs cannot, clearly, be lifted whole from general education writing-assessment procedures, more than one chair reported making thankful use of the expertise of faculty members in this area. (Writing assessment is a well-developed subfield in itself; for discussion of issues in the assessment of writing, see White; White, Lutz, and Kamusikiri.)

Another chair reported “mixed reviews on faculty response” but added, “I think [assessment] will have, in the long run, a beneficial effect on faculty members’ consciousness about curricular and departmental issues. It will place professors’ work more clearly in the context of student learning—not a bad development, I believe. Meanwhile, since we aren’t sure where we’re headed (and we feel whatever we do will be a lot of work), there’s also some cautiousness.” Yet another noted that in a department where the faculty was generally “hostile” in the beginning, most “have found [the] assessment program to be not only beneficial to the program but to them as individual professors. [The department’s] major requirements have been revised as a result of assessment, and a number of changes have been made in the way in which particular courses are taught.”

Chairs of departments where assessment was still prospective or in the early stages saw faculty members’ skepticism as reflecting lack of knowledge, mistrust of the

motives behind the mandates, and frustration with the short time allowed for implementation and the consequent difficulty of formulating plans appropriate to their programs and astute enough to elicit useful and pertinent information. In the words of one department chair, “[department members’ views] range from skepticism about the state’s motives in mandating assessment to resentment about the time and intrusion into academic freedom represented by the program to a positive feeling that the process helps us to clarify our goals and learn more about our students.” Whether the mandate for assessment of student learning was a “dust storm on the horizon” or a new and major responsibility in the chair’s office, faculty members often worried that they would receive insufficient resources to do the job well or to keep it going once engaged.

Those in departments where the process was well under way sounded a different note. Most had found their faculties could turn assessment into an opportunity to gain useful information about their programs and bring faculty members together with a new awareness of needed program revisions. “When this started we were all outraged,” one chair reported, “[but] faculty members from various departments are now talking with each other about goals and teaching strategies in ways that I think are productive. Assessment becomes a means and not an end in itself,” helping the department clarify how general education courses and upper-division courses for English majors relate to one another and how well specific department courses meet general department goals. Assessment has long been part of individual teachers’ practice in their own courses. As one correspondent observed, “Obviously, good teachers assess all the time. They assess how their students are doing, and they assess how they are doing as teachers.” In some departments, practices developed by faculty members are being applied to assessment undertaken as part of periodic program review.

But precisely because assessment may only replicate what departments already know from doing self-study and reviewing faculty members’ teaching, some correspondents doubted whether the mandates made any difference: “The results have shown us some interesting things, but altogether the results have not been surprising, so it is hard to say we have learned anything significant that we did not know or expect before the assessment went into effect. We have made changes in the curriculum and policies, but we probably would have made them anyway.” However, one department head reported that assessment led to the creation of departmental goals for the major and made the department aware of the need for a new introduction to the major course. Another department happily did away with a standardized test that had been but no longer was mandated, but it implemented its own voluntary assessment measure—holistically scored samples of student writing—and reported finding the

practice “extremely valuable”: scoring sessions “foster community, clarify standards, [enable faculty members to] share teaching ideas, provide teaching materials, and . . . significantly improve the level of instruction.”

For some chairs the gains from assessment were less a matter of new internal professional insights than a matter of new expertise in documenting excellence for audiences outside the institution: increased public skepticism leads to a new effort at improving public relations. One chair mentioned that a university advertising brochure had publicized the institution’s assessment efforts and the accreditation team’s praise for its successes. Correspondents also recognized that as tuition payers shoulder an increasing portion of the real costs of higher education and as higher education competes in state budgets with rising costs for corrections and for health and welfare, calls for accountability will only increase. They noted too that in calling for assessment the constituencies of educational institutions have motives that it is not accurate to dismiss as simple-minded concern for the bottom line.

Assessment indeed takes place in an environment of national economic trends and professional goal setting, but it also inevitably invokes a host of intensely individual institutional concerns—refining relations to original ethnic or religious institutional affiliations, raising issues involving the specific urban or regional populations the institution serves, or prompting attention to unique excellences often fostered over several generations. The most convincing assessment will be generated where these well-understood institutional particularities ground the discussion of economic realities and professional aspirations.

### Selecting and Using Evidence in Assessment

Many in ADE member departments have indicated a wish to hear something of what colleagues around the country are doing. In the expectation that chairs and faculty members will find the information useful, the committee here reports on approaches to assessment department chairs described in their letters to the committee.

When a department is called on to provide an assessment of its programs and the effect of its curriculum and teaching on student learning, it is easy to get caught up at first in questions of methods and instruments. Especially if faced with a demand to implement an assessment procedure quickly, a department can find it tempting to reach for the first instrument and evidence that come to mind, such as grades collected from a transcript review or scores provided by a commercial test. The ad hoc committee strongly recommends a more deliberative approach, one that begins with a department’s questions about how courses and programs are serving students.

Beginning with these questions lets choices about methods and instruments come into focus as means rather than ends. This approach frees a department to give proper care to each step in the procedure—identifying the goals or outcomes that are to be assessed, specifying the evidence that would be most suitable to use, gathering and interpreting the evidence, and finally reporting on findings in forms appropriate to various audiences inside and beyond the institution.

Where assessment mandates have been imposed, all but a few of the committee’s correspondents report that decisions about the form assessment is to take have been left entirely to individual departments. As a result, and also because the specifics of mandates vary from institution to institution, correspondents’ letters reveal considerable variety in departmental practices and approaches. Departments, especially those that have been doing assessment for several years, report trying and discarding several different approaches, as time reveals the possibilities, limitations, and problems of each. In the few cases where specific measures were part of the assessment mandate, the measure most commonly called for was a nationally normed test. But several departments chose to use such tests of their own accord, for reasons of cost, efficiency of administration (especially scoring), or lack of time to investigate and implement alternatives.

Other measures correspondents report using include portfolios and a senior seminar or capstone course structured to emphasize inquiry into the character and quality of learning among majors—considered both as a group and as individuals. Departments also surveyed graduates, employers, and current students; looked at student success in obtaining employment or admission to graduate and professional school; held focus groups of graduating seniors; conducted exit interviews; administered exit questionnaires and satisfaction surveys; and surveyed faculty members who taught seniors. Some departments used as reference points entering students’ American College Testing or Scholastic Achievement Test scores or, later in students’ progress, scores on common writing or rising junior exams. Some used national or locally created standardized pre- and postcourse tests on literature. Other measures included transcript analysis using an automated degree check program, statistical portraits of students entering the program (provided by their campus office of assessment), compilation of awards the department’s students received, and examination of undergraduate honors theses. Also mentioned were departmental review of syllabi for multisection courses to ensure consistency and review of course syllabi and handouts in the light of stated course goals and objectives.

Portfolios were central to many departments’ assessment plans—but many correspondents offered caveats about the work involved in scoring them or the difficulty

of adapting them to contexts where, for whatever reason, quantification is advantageous. The term *portfolio* is also subject to varying definition. For some it refers to a collection of representative writing by students, often with faculty members' annotations and an essay by the student analyzing the portfolio. For others a portfolio might also include entrance scores, a sample of freshman writing, results of a standardized test such as the Graduate Record Exam or National Teachers Exam or of a locally developed test on literature, a tabulation of courses taken and grades received in the major, letters of employment or admission to graduate school, and postgraduate writing samples and surveys. If thoughtfully conceived, portfolios offer the distinctive advantage of representing actual work students have done in response to course or program requirements. Because portfolios can represent effort and performance over extended arcs of time and across a wide array of questions, materials, and occasions, they counter the limitations of the single test or writing exercise, which affords a narrow and often misleading snapshot of students' capabilities and performance.

A few departments proposed using outside reviewers. With careful planning this procedure can provide helpful, credible information about program performance and possible future directions. It has a number of disadvantages, however, especially if assessment is to be ongoing. Administrators and governing bodies may regard outside reviewers' reports as forms of special pleading. And faculty members may complain that outside reviewers do not spend enough time on the campus or in reviewing the department materials to arrive at any useful conclusions. Using outside reviewers is also expensive.

This report does not specify in detail how individual departments should select evidence or which evidence they should select—in the committee's view these determinations necessarily depend on local programs' aims and circumstances. We do wish to offer some observations about the varieties of evidence generally available to departments. There exist three broad groups of evidence about student performance that might justify or assess the effectiveness of English department programs: institutional records concerning students, measures of student achievement, and measures of student perception of the program. Some instruments for gathering these data are available commercially, some can be derived from existing institutional records, and some must be generated by the department. Departments should remember that there is a strong history of good assessment practice in written composition and that there may be available expertise in assessment among the composition faculty that can be used to assist faculty members concerned with other programs. The principle of "multiple sources" is also key; no single measure or source of evidence is adequate to address the questions assessment poses.

Departments need to consider the options carefully before deciding whether to use a commercially available test or questionnaire or to develop their own. The advantages of the commercial test are that it can be acquired quickly and easily and that it can provide reliably scored comparisons with other institutions or with national or regional norms. The disadvantage is that what is measured may not match the department's program or goals and may not capture the department's particular strengths. A locally developed measure should certainly fit the department, but it takes time and energy to create, administer, and score reliably and may not allow the institution to compare a department's strengths with those of departments in other institutions.

Some correspondents from departments with large numbers of students in their programs noted that, for purposes of program assessment, data need to be gathered for only a sample of students assembled at random. A department might assemble a sample of students at the beginning of their freshman year, for example, or perhaps of students who declare themselves English majors in a given semester. Analysis of such cohorts is complicated, however, because the realities of students' lives mean that the composition of cohorts is almost never stable: students drop out of the educational system, temporarily or permanently; transfer to other majors or institutions, often for reasons unrelated or only tangentially related to their academic experience; or end up receiving degrees in different years. Such complications, of course, apply as well to institutions where the whole population of students in a program is assessed, and all departments should take care in analysis and interpretation.

The specific options under each category are listed below. Each option has been used by at least one department whose chair wrote to the committee. Each has a focus that might be important to a given department, each has some benefits, and each brings with it some caveats. As the following paragraphs indicate, departments can use existing data from institutional or departmental records to address many of the questions for which assessment seeks answers.

#### *Institutional Records*

Institutions and departments typically keep track of course enrollments, student retention, graduation rates, and the like, and these data have an important place in any department's array of assessment activities and approaches. If there is relative stability in the methods of collecting and presenting such data, it is possible to capture trends over time, which can be more revealing than the picture derived from data for any single year or semester.

*Course enrollments* can be used for a variety of purposes. If examined over time, such data can show enrollment

trends in specific courses. Departments can also analyze them by groups to see which sorts of students are taking which sorts of courses and at what point in their careers. Comparisons can also be made among departments if such are warranted. The figures can show relative participation in facets of a program, but their interpretation is often confounded by a variety of nonmaterial factors such as scheduling.

*Course-sequence audits* can reveal typical itineraries students take as they progress through the major or some other program. Alone, they say little about the performance of students; combined with grade-point averages, they might give some index of the stringency of the program.

*Graduation and placement data*—graduation rates and rates of placement into employment and graduate school—can, when available, serve as evidence of program performance. These figures may be useful, but they can be shaped by the market. Job-placement data can be unreliable because the first job may not be in the graduate's future career.

*Grade point averages* can be gathered for the particular English program as a whole and measured against averages for comparable programs at the institution or prior cohorts of students. Averages can also be derived for specific courses or categories of courses as well as for particular groups of students. Using averages is an easy means of showing overall achievement, but it may be difficult to infer the quality of the program from the performance of the students, whose level of achievement might stem from other causes. The results may bring charges of grade inflation from members of the department or from outside, so that they can be a two-edged sword.

*Course pass-fail rates* would normally be reported as percentages and can be aggregated by major or level of course to provide evidence of the stringency of courses. They are subject to the same problems as grade point averages are.

*Awards, student publications, and prizes* show the annual and cumulative record of a department's students in such prize competitions as scholarship awards, honor societies, literary magazines, and who's who directories. These provide a good measure of the peak moments of a program and its students, but the sample is often small.

### *Student Achievement Measures*

Institutional records like those mentioned above give information about students' progress through courses and programs. But most pointedly, assessment raises questions about what and how well students actually learn. A variety of measures of student achievement can help answer these questions.

*External tests* are available from a variety of vendors, including the American College Testing Program (ACT)

and the Educational Testing Service (ETS). Many of them are adaptations of tests developed for mass testing of large populations; some can be scored by the department, but some need to be sent to the vendor for scoring. The respondents mentioned tests including the Literature in English test and the rising junior exam from ACT, the Major Field Exam and the Graduate Record Examination from ETS, and the National Teachers Exam (for English education majors). Most of these tests allow for comparison of the students with a national normative sample, which can be a distinct advantage for an institution. One problem with such tests is that they may measure a general level of reading or language performance and may not be specific to the program that the student has taken. Another problem is that performance on these tests is best predicted by performance on earlier and similar tests, such as the Scholastic Achievement Test, since both kinds of test are designed to differentiate among students rather than to show the relative or absolute learning in a subject brought about by a program or course.

*Exit or rising tests* (given at the end of a program or as a qualification to enter one) are tests prepared and scored by the department, the institution, a commercial vendor, or the state to ascertain student proficiency in specific areas (e.g., writing). In writing, in particular, there is a good history of careful practice in the development and marking of these sorts of tests using holistic or primary-trait scoring. There are, however, limitations to assessing change brought about by particular courses through pre- and postcourse tests. Measurement experts have found that so-called gain scores pose a number of technical problems and should be viewed with skepticism.

*Major examinations*, often oral examinations or interviews, or alternatively essay examinations, can be prepared by a department and administered to all majors. Such tests can be well designed to reflect the department's goals. However, departments should take care to ensure reliable grading by the members assigned to judge the tests. Judges need to use a common scoring rubric, and it is best to have multiple judges to check the agreement of the scores given students. If the examinations are oral, there should be some means of ensuring that comparable questions are asked.

*Theses* can be required of all majors. Grading is usually performed by a committee of faculty members in the department or by an external team. Theses can provide a good example of some of the best work of students, but grading them is time-consuming, particularly for a department with a large number of majors. Departments also need to take care that the standards for grading are comparable across topics and types of theses.

*Capstone courses* combine some of the elements of a major examination and a thesis. It appears that such courses are being used more frequently than previously,

and they provide a variety of forms of assessment of the program, from oral examinations and reviews of papers, independent research projects, or portfolios to interviews concerning the quality of the student's experience in the program. The work in these courses may be collaborative and thus may allow for the students to provide some of the assessment or to serve as focus groups for interviews and surveys. The syllabi themselves can be used for program assessment, as they indicate the kinds of work the faculty considers students able to do. The advantages and disadvantages of the capstone course are similar to those of theses and major examinations.

*Portfolios* may contain a variety of papers, logs, and examinations prepared by students, as well as an autobiographical statement and an explanation and defense of the contents. As noted above, the term *portfolio* is subject to varying definition. The problem of scoring portfolios for purposes other than assessment of the individual student is similar to that of scoring other types of student work; in addition, the scoring can take a long time, and it can be difficult to use portfolio scores to quantify student performance if the department wants or needs to. One major issue in the use of portfolios (and other examinations and papers) is the possible legal limitations on their use. Since students hold rights of intellectual property over portfolios and their contents, they control how far their writings may be distributed. Reports on student performance need to be handled with discretion, and a department needs to foresee who will be entitled to view the work of students and what sort of permissions need to be obtained. Therefore it is important to be clear about what will be reported only in aggregate—how *groups* of students did, with the anonymity of individuals preserved—and where individuals' work will be presented or referred to by way of example.

*Inventories of reading and writing* include students' lists of what they have read for their courses and as outside reading as well as lists of the kinds of writing they have done. Such lists serve as a supplement to or in lieu of a portfolio. As with many other measures, it is often hard to determine whether the results emerge from the program or from the students.

*Classroom videos*, particularly of senior seminars, can give a sample of the level of discussion that goes on. Some students may prepare videos of their own work. These videos can effectively show a variety of work, especially collaborative work, but the technical quality must be sufficient to impress an outside audience.

### *Student Perceptions*

Most programs routinely gather information about student perceptions and satisfaction. Such data can round out and complement the evidence provided by in-

stitutional records and student achievement measures. Many of the strategies for uncovering student perceptions have the additional advantage of making students more active partners in shaping their educational experience.

*Course ratings* based on questionnaires distributed at the end of a course are often provided by a department or an agency of the campus as a normal part of faculty and course evaluation. Data may come from machine-scored questions, open-ended questions, or both. The questionnaires are often geared to large classes, and the results can be biased toward easy courses. But if properly designed and thoughtfully read, student course evaluations can be a credible and important source of information (see Robert Boice, "Countering Common Misbeliefs about Student Evaluations of Teaching," reprinted as an appendix in Elbow 6–7).

*Exit interviews and questionnaires* are often a part of the normal program evaluation of the institution. They are usually given to seniors in the last semester; often they contain questions about many facets of campus life, and departments may have to request additional questions pertaining to their programs. Capstone courses can function as at least one locus for exit interviews conducted by the department of its majors. The questions and the way in which the interview is conducted need to be carefully scrutinized. It is usual to have an interview schedule prepared so that all the interviewees will be asked essentially the same question. Interviewers need to be careful not to lead the students toward a desired answer.

*Focus-group interviews* are generally conducted with a specially selected group, such as prospective and current students and recent graduates. The group is interviewed collectively and the results are transcribed. Again, one class hour of a capstone course or more could be adapted as a focus group. However, the sample can be biased, and what criteria the students are using needs to be carefully explored.

*Alumni interviews* are often conducted by university or alumni officials. Departments may ask for additional questions or may conduct their own interviews and construct and distribute their own questionnaires. The interviews give a longer-term retrospective view of a department, but they are difficult to arrange and conduct, and they may prove to be out-of-date with respect to the current goals and programs of a department.

Each of the three types of evidence has advantages and drawbacks. Student records provide a history of student performance, but it is hard to interpret grades and enrollment data in the light of written norms or standards or data from other institutions. A grade masks the criteria that went into assigning it, which can include matters such as attendance and participation as well as academic performance. Student achievement scores provide

information about performance according to a criterion or according to group norms, but such data are plagued by issues of reliability and validity—whether the score reflects a true judgment of performance and whether the score reflects the curriculum of the department. Student questionnaires provide valuable information on the satisfaction of the clients of a department’s program, but they do not always inform an observer of the content or quality of learning.

The committee does not recommend or discourage use of any specific measure. It does, however, urge departments to consider all options carefully and recommends that departments use a combination of kinds of evidence to provide an efficient and effective assessment of student achievement in the light of various institutional constraints. Departments should exercise due care in the design, scoring, and interpretation of each of the assessment instruments that it selects. The committee also recommends that departments use campus experts to help design and implement the assessment, so that the measures may be suitable to the particular institution and so that local faculty members are in charge.

#### Hints, Tips, Words to the Wise, and Other Cautionary Observations for Department Chairs

*Read the mandate.* Begin by reading the mandate to which you are being asked to respond—and read carefully to see what you are and are not being asked to do. You will often find that the mandate allows you considerable leeway in determining both the focus and the methods of assessment. Don’t jump to dire conclusions. The job at hand will usually turn out to be smaller than you think, less of an imposition, and more in line with your department’s goals in thinking about and evaluating students.

Like many of our correspondents, you may find that the mandate is vague. At first you may be inclined to find the vagueness frustrating. But vagueness can be good: it is an invitation for your department to set the terms of the discussion. As you do so, it is a good idea to check in with the administration—the people who received the mandate and passed it on to you—to let them know how you and your colleagues think you can best define the terms and goals of assessment.

*Begin with your department’s own goals and values.* Assessment can serve a department by getting people talking about the goals and values of education in English. It is crucial to be sure that your assessment procedures are pointed toward the areas of learning your department believes are most important. There is often a temptation to describe outcomes in broad, general terms (like “critical thinking” or “historical understanding”). It makes more sense, because it yields more-useful results, for the faculty

to name as specifically as possible the range of things its members teach students, both majors and nonmajors: literary history? the writing of the critical essay? library research? great books and great ideas? literary production in cultural context? experimental prose forms? (The purposes will almost certainly be multiple.) It is important to take into consideration your department’s mission and identity, both locally and in relation to the profession. What counts as important depends on your department’s programs and goals, and these should determine the focus of assessment. If your goal is to demonstrate that your department is successful at what it does, your assessment plan will differ from what it will be if you intend to highlight problems and motivate change.

*Be circumspect.* At the beginning and at the end of formulating an assessment plan—that is, when you start thinking about your department’s educational goals and when you start reporting on whatever data you’ve collected—be careful not to make claims you can’t support and promises you don’t want to be held to (or that might later be used against you). The English department makes contributions to broad institutional goals, but take care that your department isn’t held solely responsible for meeting those goals. A single course is not by itself going to solve students’ literacy problems. The English department isn’t the only place where students learn how to think, write, or read critically. Reading literature isn’t a foolproof way to produce democratic citizens. By making grand claims you can set yourself up for big trouble down the line.

*You don’t have to start from scratch.* Instead of assuming that assessment must involve adding an entirely new apparatus to your department’s infrastructure, think of building it into or onto existing practices. And look in your files—you may already have information you can use to document student achievement. Your department may already have a capstone course, a senior portfolio or project, a large-scale examination of student writing, or a comprehensive exam that could serve as an ideal context for assessing student progress. Assessment might also be woven into an advising process or linked with the activities of an English club or honor society. Many departments already conduct some kind of alumni survey or follow up with graduates, providing an opportunity for questions about the program’s effectiveness. Building aspects of assessment into multiple existing practices raises the odds of gathering significant, useful information without adding a layer of bureaucracy.

Remember—look first at where and how you already review the work of students. Then you can think about where you might find or solicit useful information and how to interpret and present your data. A regular meeting of those who teach the senior seminar might constitute an assessment of majors. Collecting sample portfolios

from sophomore literature classes could provide the database for a larger assessment of general education.

*More is better (but do more sparingly).* If you want to construct a comprehensive account of program effectiveness, no single type or source of evidence will suffice—it is important to use multiple measures. But the need for multiple measures and methods must be balanced by thoughtful concern for cost and for unintended consequences such as the potential for intrusiveness that detracts from the quality of instruction and learning.

The forms of assessment many in the humanities find the most attractive are also often the most expensive. In particular, don't forget the hidden costs of faculty members' time and goodwill in (to take an instance) portfolio assessment, one of the most attractive forms of assessment in English. Students can also pay a price. In the day-to-day life of classrooms, where grades are so central and often so distracting, more testing and measuring can inadvertently leave students feeling intimidated, frightened, attacked, or pigeonholed. The Hippocratic rule applies: first, do no harm.

*Take due account of student motivation.* If assessment does not occur within existing contexts (such as courses, required projects, and the like), be sure to think about how and why students will be motivated to participate in whatever assessment activity you are contemplating. Many campuses that have devised special surveys and self-standing tests in response to an assessment requirement have found that students' participation and motivation are low, especially among seniors only days away from graduation. As a consequence, departments are left with data that misrepresent the effectiveness of their programs' curricula and teaching. Some have addressed this problem by making completion of the assessment instrument a degree requirement. Involving students as partners in the design and implementation of assessment activities has also helped solve the motivation problem in some settings—but such an approach probably means repeating the exercise each year as students graduate.

*Don't forget to consider the legal and ethical implications of whatever you plan to do.* Assessment activities raise concerns about privacy and the use of human subjects, particularly when assessment involves using student work (papers, projects, interviews, exams, portfolios, and so forth) for purposes beyond those of the original course for which the work was prepared. Most institutions have policies governing the use of human subjects and access to student records. Laws governing intellectual property also limit your ability to reproduce or publish student writing. Assessment puts into the spotlight the question Who owns student work? And the answers are not always clear. Take due account of students' rights, including their legal rights to their intellectual property. There is probably someone on your campus, in the psychology

department, the legal office, the writing program, or an office of measurement and evaluation, who can advise you. It is certainly crucial to have these conversations early in the process of planning assessment. The safe course is to seek written agreement, in advance, from all relevant parties about what data is gathered for what purposes and for whose eyes.

*Be wary and alert.* You also should think carefully about who will have access to the information you gather, and how the information will be used, stored, and reported. You may not want what you report to be tied to particular courses, faculty members, or even programs. It may not be valid or appropriate to compare classes, years, or campuses. How to understand and report what you collect involves problems of validity and reliability that are technical and often tricky. Look for people on your campus who have expertise to guide you through these problems.

Assessment has acquired its own jargon; terms mean different things to different parties, and these meanings have significant implications, especially if you are caught unawares. A term that caught on early in assessment circles is the phrase *value added*, which refers to gains in student learning that can be attributed to courses and programs. The thrust of the term was to enforce the general notion that the quality of an educational program should be understood as a matter of how much students learn (rather than how smart they are when they enter or how plentiful the resources of the program are).

Some, however, use the term *value added* to refer specifically to gain scores calculated from pre- and post-tests. When the term is linked in this way to specific measures that are supposed actually to capture value added in a single number, all kinds of confounding technical and reporting issues arise. If your seniors perform better than your freshmen, is that because of some aspect of the program, because of something that happened somewhere else, or perhaps simply because they matured? Or if your seniors don't perform as you had hoped, their performance might be for reasons outside your control (they didn't study, they blew off the exam, or they were working thirty hours a week). Be wary when people start talking about value added. It is not a simple matter.

*Is it useful?* To be useful, the information derived from assessment must be connected to issues or questions people really care about. Assessment approaches should be designed to produce evidence that relevant parties will find credible, suggestive, and applicable to decisions that need to be made. The point of assessment is not to gather data and return "results"; assessment is a process that starts with the questions of decision makers, involves them in the gathering and interpreting of data, and informs and guides continuous improvement.

A corollary is that departments will do well to think in advance about who, exactly, wants to know and will use

the results they report. Standardized tests are tempting because they are quickly obtained and easy to administer. But few departments find the scores they provide credible, interesting, or useful, and many who use them eventually abandon them. When programs instead take the time to design approaches that inform discussions about issues actually on people's minds, they are in a better position to develop assessment procedures that are more productive and efficient in the longer run.

*Be realistic.* Many of the most appealing methods of assessment—portfolio review, interviews, observations—are also the most time-consuming. You can make scarce resources, especially faculty members' time, go further by using sampling. For purposes of assessment, departments with large numbers of students don't need, for example, to examine every senior major to derive meaningful information about the senior class. What's needed is for departments to look at a reasonable, representative random sample—perhaps between ten and twenty percent for a large group. Depending on the questions being asked, you may want to oversample certain categories of students. For instance, if you are trying to learn more about minority students, you may want to include more than twenty percent of the minority seniors in your sample.

You don't need to do everything at once. In fact, doing a few modest things one after another over time may be the most productive approach.

*Who should be at the table?* The particulars of mandates and methods aside, assessment is best approached as a process of bringing stakeholders together to discuss and negotiate their common enterprise. In an academic department the stakeholders include faculty members, certainly, but also students; others outside the department (employers for whom your graduates work or faculty members from graduate programs they go on to attend, for example) may provide additional important perspectives on a department's goals and objectives. The more people who are invested in the process, the more likely it will have broad support (particularly in a climate of suspicion) and effectiveness. Viewed from this angle, assessment is an opportunity to have a kind of conversation that we seldom have in a purposeful way, a conversation about what we do in undergraduate education and why.

*Don't blow it off.* Like it or not, assessment is with us now and will be for the foreseeable future—built into state-level mandates and into the accreditation process in every region of the country. It is clear from the reports we received that it is not easy for faculty members to embrace the idea of assessment. What is also clear is that we are at a moment when faculties of English departments need to find ways of explaining to the public who they are and what they do. Moreover, the moment of assessment has proved to be an opportunity for departmental conversation and reflection that are otherwise too rare.

As one chair reported, "Our faculty lost a great chance when it failed to take this mandate seriously."

*Don't lose sight of the big picture.* Remember that requests for assessment are directed at institutions as well as departments. They can thus become occasions for potentially useful institutional as well as departmental change. Assessment asks faculties to look at what they and their programs do, to note strengths, and to consider areas in which changes may be called for. Assessment mandates remind us that learning can and should be talked about at the program level as well as at the course level and by the faculty as a collectivity as well as faculty members as individuals. Approached as an activity faculties engage in as a group, assessment can put important issues on the table.

The committee heard two clear messages from department chairs who have experience with assessment: (a) assessment need not be as difficult or overwhelming as it at first glance appears, and (b) doing assessment offers a chance for colleagues to sit down and have a kind of discussion they do not usually have—about what students are learning, know, and can do in relation to the faculty's teaching and the department's curriculum.

Behind the advice we offer are some assumptions about good practice. Although a department may need to generate evidence to answer real questions about undergraduate education for various and even skeptical audiences, the primary purpose of assessment is to improve what we do for students. We are not giving tests for the sake of giving tests or simply to satisfy bureaucratic requirements. Our first responsibility is to our students, our colleagues, and the values of our departments and our discipline.

Prepared by the ADE Ad Hoc Committee on Assessment

March 1996  
Association of Departments of English  
10 Astor Place  
New York, NY 10003-6981

---

## Works Cited

- Atwell, Robert H. "Putting Our House in Order." *Academe* July-Aug. 1994: 9-12.
- Benjamin, Ernst. "The Decline in Academic Autonomy in Higher Education." *Academe* July-Aug. 1994: 34-36.
- Elbow, Peter. "Making Better Use of Student Evaluations of Teachers." *ADE Bulletin* 101 (Spring 1992): 2-8.
- Hutchings, Pat, and Ted Marchese. "Watching Assessment: Questions, Stories, Prospects." *Change* Sept.-Oct. 1990: 12-38.
- Longanecker, David A. "The New Federal Focus on Accreditation." *Academe* July-Aug. 1994: 13-17.

- 1996 *Higher Education Directory*. Ed. Mary Pat Rodenhouse. Constance Healy Torregrosa, ed. emeritus. Falls Church: Higher Education Publications, 1996.
- "NPB Recommends Accreditation Standards." *Higher Education and National Affairs* 26 Sept. 1994: 5.
- Poston, Lawrence. "Statehouse, Countinghouse, and Our House." *ADE Bulletin* 114 (1996): 14–18.
- Warren, David L. "Why Faculty Should Care about Federal Regulation of Higher Education." *Academe* July-Aug. 1994: 18–19.
- White, Edward M. *Teaching and Assessing Writing*. San Francisco: Jossey-Bass, 1985.
- White, Edward M., William Lutz, and Sandra Kamusikiri. *Assessment of Writing: Politics, Policies, and Practices*. New York: MLA, forthcoming.
- Banta, Trudy W., et al. *Assessment in Practice: Putting Principles to Work on College Campuses*. San Francisco: Jossey-Bass, 1996.
- Conference on College Composition and Communication. Committee on Assessment. "Writing Assessment: A Position Statement." *CCC* 46 (1995): 430–37.
- ERIC Clearinghouse on Assessment and Evaluation home page. Online. World Wide Web. Available [http://www.cua.edu/www/eric\\_ae](http://www.cua.edu/www/eric_ae).
- Gould, Christopher. "Assessing Teaching Effectiveness in English: Procedures, Issues, Strategies." *ADE Bulletin* 102 (1992): 44–52.
- Hansen, Richard E. "Developing an Assessment Process: One Department's Story." *ADE Bulletin* 113 (1996): 46–49.
- Harris, Charles B. "Mandated Testing and the Postsecondary English Department." *ADE Bulletin* 107 (1993): 4–13.
- Hart, Dominick J. "Assessment and Accountability." *ADE Bulletin* 113 (1996): 42–45.
- Martin, Carol. "Information and Education: A Fable for Assessment." *ADE Bulletin* 113 (1996): 39–41.
- McNew, Janet, and Cindy Malone. "Assessing the English Major: A Case Study." *ADE Bulletin* 107 (1994): 28–30.
- Terenzini, Patrick. "Assessment: What It Is and What It Isn't." *ADE Bulletin* 107 (1994): 14–17.

---

### Selected Resources

---

AAHE Assessment Forum home page. Online. World Wide Web. Available [http://www2.ido.gmu.edu/AAHE/Special\\_Programs/Assessment.html](http://www2.ido.gmu.edu/AAHE/Special_Programs/Assessment.html).