

Assessing the Foreign Language Major at the State University of New York: An Interim Report

Catherine Porter Lewis

WHILE the concept and practice of assessment in education must be as venerable as education itself, the term *assessment* has in recent years become something of an academic buzzword. In many institutional contexts today, to assess means to conduct a comprehensive evaluation of a broadly defined object or target group. Unlike classroom instructors, who tend to be primarily concerned with discrete measures (such as test grades) for evaluating the performance of individual students, assessors seek to accumulate multiple measures (including, but not limited to, existing "archival" data, standardized tests, specially constructed test instruments, and interviews and questionnaires) in order to describe broad outcomes for groups of students—for example, changes in cognitive knowledge, skills, attitudes and values, or behavior (Ewell). In a well-designed assessment project, the assessors determine precisely what they are seeking to evaluate and then adopt or devise a set of instruments, use these to gather data, subject the data to analysis, and report their findings (Adelman).

If there is some consensus among professionals about the what and the how of assessment, answers to the why are more varied. Assessment for whose benefit? for what purpose(s)? "To improve institutional effectiveness" is an often repeated answer that may merely beg the question of what objectives are to be met through improved effectiveness. Who sets these objectives, who evaluates them, and who determines how success in meeting them is to be measured? How are the data that are generated to be used? Will the results of an assessment project influence funding or personnel decisions? Faculty members asked to participate in an assessment project rightly raise such questions and inevitably wonder about hidden agendas.

In October 1985, when the first National Conference on Assessment in Higher Education was held in Columbia, South Carolina, some influential and alarming national reports on higher education and its

shortcomings had begun to focus faculty attention on student outcomes and how they were (or were not but might be) measured (Association; Study Group). The legislatures of several states seemed inclined not to wait for faculty response but instead to rush to impose top-down measures for assessing the quality of state institutions of higher education (Marchese 3). It was not difficult to detect an underlying self-protective agenda in some of the discourse at the conference (and in other academic discussions of assessment), an unspoken appeal that might be formulated reductively as follows: "Let's figure out what assessment is and *do* it, before 'they' step in and discover that we're not measuring up."

One enterprising response to the call for local initiatives came from E. Thomas Moran, vice president for academic affairs at the State University of New York (SUNY), Plattsburgh. In the spring of 1987, Moran submitted a grant proposal entitled "Comprehensive Assessment in Disciplines" to the Fund for the Improvement of Post-Secondary Education. The project called for two representatives from each of five departments at five SUNY campuses to meet two or three times a year for three years, in order to develop assessment strategies and instruments that would be compatible with the curricular aims of each department. The project was made collective in an attempt to strike a balance between national norm-referenced testing at one extreme and idiosyncratic local measures at the other. The proposal was funded, and the five disciplines that were selected (essentially on a volunteer basis) included foreign languages.

In April 1988, at the initial training workshop con-

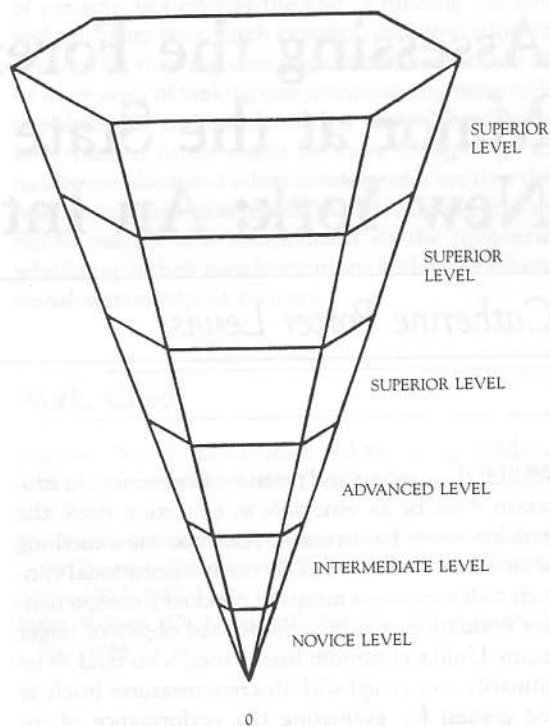
The author is Chair of the Department of International Communications and Culture at the State University College of New York, Cortland. This article is based on her presentation at ADFL Seminar East, 1-3 June 1989, in Athens, Georgia.

ducted under the grant, the participants became aware that the effort to seize the assessment initiative was indeed one aspect of the project's agenda. (As a colleague translated this part of the message later, "If we're driving the bus, we're not so likely to get run over by it.") In his charge to workshop participants, however, Moran placed much greater emphasis on the goals recommended in the *Involvement in Learning* report (Study Group): to increase student responsibility for the educational process; to define and make explicit appropriate expectations for learning; to provide useful and timely feedback; to supply a framework for ongoing evaluation and improvement of advising, teaching, and curricula.

Moran's introductory remarks thus set a strongly positive tone and posed a daunting challenge. The participants from each field were invited to consider all aspects of their own discipline, to review its basic assumptions about knowledge and learning, to delineate its accumulated body of knowledge, and to determine their own perspective on its value, meaning, and importance; after reaching consensus on these issues, they were to devise and implement a comprehensive assessment plan consistent with the discipline and their own institutional framework.

In the foreign language working group,¹ we decided as a preliminary move to restrict our scope to the liberal arts language major, excluding related programs such as those leading to secondary certification or to a degree in international studies. We also ruled out such noncognitive concerns as students' affective behavior and their ability to make ethical decisions and value judgments. We judged that these traits are more appropriately examined in an assessment of general education or of the overall college experience. We decided to include critical-thinking skills within our purview, however, on the grounds that an increase in the complexity of intellectual activity is inherent in the ascending scale that would be used for assessment and evaluation in each area.

The scale in question is the familiar inverted pyramid (Medley 18; see accompanying illustration), representing a continuum from the absolute novice to the mythical "educated native speaker" (or the hypothetical "trained disciplinary specialist," where the focus is on content); we agreed to try to use the ACTFLETS proficiency scales as the organizing principle that would inform the entire assessment effort. With the ACTFL oral-proficiency interview as our instrument for assessing speaking skills, our task was then to adopt, adapt, or create comparable scales and instruments for entry- and exit-level assessments of each of the other skill areas—listening, reading, and writing—as well as of the two primary content areas we decided to address initially, literature and culture.



Graphic representation of the ACTFL rating scale

When we turned our collective attention to literature, we found ourselves focusing initially not so much on content—no one seemed to have a vested interest in a particular vision or revision of the canon—as on the skills our students bring with them as readers and on the skills we hope they will acquire during their undergraduate years. We chose to define the novice not as a perfectly blank slate but as a college freshman with as much training in literature as is offered in the standard high school English curriculum. From that starting point, Sylvie Debevec Henning produced a proficiency scale in draft form that was revised and refined as group members used it to assign holistic ratings to writing samples we collected during 1988–89.² The French subgroup selected a short story as the basis for an assessment exercise administered to groups of entry-level (that is, fifth-semester) and more advanced French students. Given the text to prepare overnight, candidates are asked to come back the next day and write for an hour, either in the target language or in English, discussing the text “as a work of literature,” with whatever meaning they attach to that phrase. Our ratings of some two dozen samples turned out to be quite consistent, although they led us to revise some aspects of the rating scale itself.

Our approach to culture and civilization is based on a similar scale, adapted by Gisèle Féal.³ Like theoral-proficiency guidelines, the culture-civilization grid reflects a progression from the Novice level, where students know memorized material and isolated concrete facts but have virtually no social competence, to the top of the Superior level, where they have the knowledge, understanding, and social competence of a bicultural native. Our tools for evaluating cultural competence as a skill tentatively include a combination of interviews, role-playing exercises, cultural vignettes (with a choice of follow-up behaviors), and self-assessment devices. For assessing cultural knowledge, including knowledge about literature and literary history, we experimented with multiple-choice formats before tentatively adopting a grid-completion exercise based on a model developed for German and contributed by Jürgen Kleist.⁴ In a related project, some of us at Cortland are undertaking a survey in France to attempt to identify the elements of French history that educated native speakers view as indispensable components of their own cultural literacy.⁵

To assess writing skills, we have developed a three-part test that borrows heavily from a government scale described by Martha Herzog of the Defense Language Institute (155) and from a model proposed by Sally Magnan (111). The test was administered on several campuses in the fall of 1989, and the results will be evaluated at a subsequent meeting. The next task on our collective agenda is to develop reading and listening tests along similar lines.

In a discussion of assessment, Moran warns against assessing only what is easy to measure. Although we have not entirely avoided this pitfall, we are finding that the proficiency scale provides a conceptual framework that facilitates and legitimizes holistic judgments. To be sure, our working group has taken great care in the selection and design of test instruments; still, there is a sense in which we know that any choice will be arbitrary, any approach tentative, any results approximate. This awareness is oddly reassuring; we need not wait for the perfect assessment package in order to proceed. The process itself is of primary importance, as we have already begun to recognize in our home departments. We have come to suspect that the assessment project will have "succeeded" not when we have filled in all the blank spaces on our list of instruments or even when we have collected and analyzed interesting data but rather when and if the project promotes a substantial shift in departmental thinking, so that the foreign language major is viewed as something other than a set of courses to be "covered," a number of credits to be earned.

The two remaining elements in our assessment plan, the handbook and the portfolio, may be crucial factors in bringing about such a shift. A handbook for

majors, in versions tailored to each language and each institution, will be distributed to students as soon as they declare the major.⁶ The handbook will

1. Explain the goals of the major program; present the concept of a learning continuum; outline the skill and content areas to which the concept of a continuum will be applied; stress ongoing assessment and self-assessment
2. Summarize local requirements and present sample paths for meeting them (degree plans); provide information about study-abroad possibilities, optional minor or joint major programs, and alternative tracks within the major; provide a tracking form on which a record of credits, grades, and completed courses can be entered
3. Present the ACTFL-ETS guidelines for oral proficiency and comparable guidelines for the other skill and content areas
4. Describe the entry- and exit-level assessment instruments that will be administered
5. Present a reading checklist—a comprehensive, chronologically organized list of authors and titles—with indications about how it is to be used
6. Describe the portfolio's purposes and anticipated contents; provide a checklist for portfolio items

The tracking forms and checklists included in the handbook, regularly updated, will be the cornerstone of the student's portfolio. This portfolio, to be maintained by the student and reviewed periodically with the adviser, is intended to provide baseline measures as reference points, along with tangible evidence of progress over time. The following are required components of the portfolio:

1. Tracking forms and checklists from the handbook
2. Nonconfidential entry-level examination materials, along with a summary of all assessment results
3. At least one example of a research paper written in the target language
4. Examples of the student's best work from a variety of courses taken in the target language (if possible, including courses taken abroad)
5. At least one oral production on tape each year (for example, a reading of a poem)

The remaining components are optional:

6. Examples of creative writing in the target language
7. A personal journal written in the target language
8. Related work from other courses (in English if appropriate)
9. Videotaped material (for example, a skit in which the student has participated or an interview with a native speaker)

We envision asking the student to design an exit-level plan for continuing development as a culminating exercise for the portfolio.

Except for the material pertaining directly to assessment instruments, most of the handbook's components already exist, and many items destined for the portfolio are already likely to be generated in the course of a typical student's career. The novelty in this assessment plan lies in the degree of explicitness with which we shall now be able to outline our expectations for each student and in the concrete evidence of progress we intend to gather. The collective work we have undertaken to clarify our own answers to a question like "What is a French major?" has made us much more aware of our responsibility to delineate the discipline for our students.

While the process undeniably means extra work for faculty members, especially in the early stages, it entails some unanticipated benefits as well. The ability of colleagues from different campuses to agree on a characterization of the major that we can all more or less comfortably endorse on behalf of our institutional colleagues has provided us with helpful feedback about the coherence of our discipline and the appropriateness of our expectations. The oral-proficiency testing movement had already taught us that many assumptions held by the public at large and by the teaching profession itself about teaching for "fluency" and "mastery" have been in part unrealistic and inappropriate; by designing and implementing our own assessment program we can respond confidently to those outsiders or administrators or colleagues who fault us for not producing "fluency" in two years and full-fledged "cultural literacy" in four. By pooling our experience with objective measures of student performance over time, we expect to tailor our curricular goals and the means we use to reach them in a variety of ways (although not necessarily in the same ways from campus to campus) to provide a better fit. The data generated by our assessment instruments will be used to calculate norms and to make generic comparisons, not to foster intercampus rivalries; locally, the data will be used for curricular reform and collective goal setting, not for grading individual students or evaluating individual faculty members.

Thus, even in its embryonic state, the assessment plan outlined here is beginning to develop a not so hidden agenda of its own. Students who know clearly what is expected of them can take more responsibility for setting their own academic goals and for assessing their own progress. The assessment process brings about a shift in focus from what teachers "want" of students to what students want for themselves, encouraging a sense of empowerment in students. The conceptual and pedagogical usefulness of the ACTFL-ETS scale as an organizing principle is be-

coming increasingly clear in the context of assessment.

As students and faculty members alike come to see how the learning process forms a continuum, we anticipate that a number of subtle but critical changes will come about in attitudes and expectations. Given the model of the ascending, expanding, open-ended scale, it is obvious that no undergraduate should expect to reach the "top" by commencement day. Indeed, it should become obvious—and acceptable—to students that most faculty members are also still progressing unevenly toward unattainable ideals. Just as no starting point is unacceptably low (we are delighted to have majors whom we have nurtured from the absolute Novice state), no arbitrary end point is established as a goal to achieve or a hurdle to jump. The use of parallel scales to measure diverse skills and content areas should help to encourage self-assessment in relation to the student's own starting point and objectives rather than in relation to the progress of peers, as each student is bound to have an idiosyncratic pattern of strengths and weaknesses. Finally, the assessment process helps us keep in focus our own multiple responsibilities as faculty members; it is up to us, after all, to ensure the orderly evolution of our curricula, to enable our students to set and meet their own goals within our disciplinary context, and to model the concept of continuing growth along the various disciplinary paths we have chosen.

The results of the SUNY assessment project in foreign languages are not likely to be conveniently quantifiable or transferable. Although we will readily share our working tools and principles, we do not expect to produce a package of instruments, procedures, or data on norms that can be adopted as such by colleagues elsewhere. The focus on the process of assessment rather than on the products implies that any group embarking on a comparable project has to define its own objectives, chart its own course, and measure its own progress according to local circumstances. Our own interim report, already the third in a series,⁷ will necessarily be superseded by other attempts to assess where assessment is taking us. The most telling outcome of an assessment project, then, is perhaps not a final report but an ongoing commitment to critical scrutiny—from which the proficiency "movement" and the assessment process themselves must not be exempt.

Notes

¹The original working group, chaired by Craig Sample (French, Plattsburg), included, in French, Ruth Antosh (Fredonia), Hazel Cramer (Cortland), John Cross (Potsdam), Gisèle Féal (Buffalo), Sylvie Debevec Henning (Plattsburgh), and Catherine Porter Lewis (Cortland) and, in German, Marion Sonnen-

feld (Fredonia) and Alex Wachslar (Potsdam). Wachslar has since left the project; Jürgen Kleist (German, Plattsburgh), Robert Pontorio (French, Cortland), and Irmgard Taylor (German, Cortland) have joined.

² At the Novice High level on this scale, for example, a student would be able to summarize a plot, describe a setting, and distinguish major from minor characters, among other things. At the Advanced Plus level, the student would be able to describe the work's relation to various contexts (historical, sociopolitical, literary, philosophical), demonstrate understanding of the function of rhetorical figures, characterize narrative point of view, and so on.

³ This is a complex scale with two major sections, "Culture/Behavior" and "Civilization/Knowledge," each further subdivided into "Content" and "Skills." Here are the descriptors for the Advanced level:

Cultural/Behavior: Content. Etiquette; taboos and sensitivity; polite requests; invitations; gifts; apologies; introductions; use of telephone; purchasing and bargaining; routine banking

Cultural/Behavior: Skills. Demonstrates limited social competence; can deal with routine social situations; comprehends common social rules; does not offend but does miscommunicate; is aware of additional social patterns

Civilization/Knowledge: Content. Geographic, historical, artistic concrete facts in limited context; general current events and policies; field of personal interest

Civilization/Knowledge: Skills. Describes basic concrete historical, artistic, social phenomena; compares and contrasts; discusses a few aspects of home and foreign country

⁴ Students are given a list of names and asked to indicate for each name the corresponding historical period, the individual's field or profession, and an achievement or a characteristic for which the individual is well known. A similar grid can be prepared for dates, geographic locations, and so on.

⁵ In the first phase of this survey, French citizens at the baccalaureate level or above are given a list of names, dates, and historical events; respondents are asked to rate on a scale of 1 to 5 the importance of each item as a component of a curriculum designed to prepare Americans to teach French. A follow-up study using the grid-completion technique described in note

4 will attempt to assess how well the respondents are informed about the items listed.

⁶ The first working version of a handbook has been prepared for German by Marion Sonnenfeld of Fredonia; a French version is in preparation at Cortland.

⁷ Hazel Cramer drafted the group's first interim report in June 1988. I drew heavily on her work for an earlier version of this paper, presented at the 1989 ADFL Summer Seminar in Athens, Georgia.

Works Cited

- Adelman, Clifford. "Models of Assessment in the Disciplines." FIPSE assessment workshop. Syracuse, 16 Apr. 1988.
- Association of American Colleges, Project on Redefining the Meaning and Purpose of Baccalaureate Degrees. *Integrity in the College Curriculum: A Report to the Academic Community*. Washington: Assn. of Amer. Colleges, 1985.
- Ewell, Peter. "Assessment: Its Vocabulary, Dimensions, and Rise as an Issue." National Conference on Assessment in Higher Education. University of South Carolina. Columbia, 14 Oct. 1985.
- Herzog, Martha. "Issues in Writing Proficiency Assessment Section 1: The Government Scale." *Second Language Proficiency Assessment: Current Issues*. Ed. Pardee Lowe, Jr., and Charles W. Stansfield. Englewood Cliffs: Prentice, 1988. 149-77.
- Magnan, Sally Sieloff. "Teaching and Testing Proficiency in Writing: Skills to Transcend the Second-Language Classroom." *Omaggio* 109-36.
- Marchese, Ted. "Third Down, Ten Years to Go." *Bulletin of the American Association for Higher Education* 40.4 (1987): 3-8.
- Medley, Frank W., Jr. "Designing the Proficiency-Based Curriculum." *Omaggio* 13-40.
- Moran, E. Thomas. Keynote speech. SUNY, Cortland, Faculty Forum on Assessment. Cortland, 17 January 1986.
- Omaggio, Alice C., ed. *Proficiency, Curriculum, Articulation: The Ties That Bind*. Middlebury: Northeast Conference on the Teaching of Foreign Languages, 1985.
- Study Group on the Conditions of Excellence in American Higher Education. *Involvement in Learning: Realizing the Potential of American Higher Education*. Washington: National Institute of Education, 1984.