REVAMPS THE COMPETENCY PROCESS FOR WRITING: A CASE STUDY

Speakers: Deborah Holdstein, Governors State University, Illinois
Ines Bosworth, Educational Testing Service, Illinois

Introducer/Recorder: Lu Ming Mao, University of Minnesota

Deborah Holdstein began by describing Governors State University, a junior, senior, and graduate institution with diversified student body. The University used to have a writing competency test for prospective juniors and seniors. Accompanying this competency test were a set of grading standards and a pass-and-fail system, both of which had continually drawn criticism from test readers and scorers alike, because they were vague and not academically sound. According to this set of standards, a passing essay must (1) respond to the stated topic; (2) have a clearly stated thesis; (3) show clear, logical organization of ideas in organized, well-developed paragraphs; (4) include supporting details; (5) demonstrate one's editing ability.

Holdstein was asked to change this system. She observed that the process of revamping an assessment program was as political as it was academic. One misperception handled around a lot was that the English teachers were determined to flunk students. Holdstein recalled that they needed someone from outside, an expert with no stake, political or otherwise, in the system, to help teachers revamp the system. Ines Bosworth from Educational Testing Service was brought in as a consultant. Bosworth emphasized that as a neutral observer, she was able to get different opinions from faculty in different departments. These discussions became extremely useful because they enabled faculty to articulate their concerns about possible changes in the testing program. Out of these discussions—and the Provost's unflailing support—came the new scoring criteria, which have four major areas: focus, organization, elaboration (support) and conventions (mechanics). These are scored with a 6-point scale; 6 being superior and 1 being seriously inadequate. This scale replaced the old pass/fail scale.

Holdstein noted that the number of questions on the test was reduced from 5 to 3 (although the test time is still 60 minutes) in response to students' complaints that the number of tasks on the old test forced them to spend a lot of time reading and figuring out questions instead of actually composing. One of the three new questions reads as follows:

"Matrimony is a process by which a grocer acquired an account the florist had. What does this quote say about the transition from single to married life? Is it accurate? How so—or how not? Again, be sure to formulate a thesis with your point of view, and use specific examples to back up your points.

Both speakers noted that one of the many merits of this new competency test is that readers can more easily score each essay according to the criteria. Moreover, the new system is fairer than the old one. Under the old system, whenever there was a split in "failing" or "passing" decisions, a 3rd reader was consulted. Under the new system, each essay receives four readings, and readers do not know whether they are the 3rd or 4th reader; thus, a lot of political heat is removed. Readers also provide students with an "analytic checklist," which informs them of the criteria used, the weaknesses in their essays, and comments from the readers.

Bosworth commented that the interrater reliability of the new test is 92% (as opposed to 73% in the old test) and that more students have passed the new competency test than before. However, Holdstein pointed out that most questions in the new test tend to be too content-laden and that the scoring criteria are too heavily weighted toward content. Nevertheless, both speakers noted that the new test has proven to be far more effective than the old one and has fostered faculty collaboration.

WE DID IT AND LIVED: A STATE UNIVERSITY GOES TO EXIT TESTING

Speakers: Phyllis Liston, John Mathew, Linda Peter, Ball State University

Introducer/Recorder: Joyce Malek, University of Minnesota

In Fall 1987, Ball State University (Muncie, Indiana) implemented exit testing for writing competency as a prerequisite for graduation. The three member panel responsible for establishing the rubrics and coordinating the testing and holistic grading discussed what they learned during this first year. Participants were given hands-on experience with the exam by writing briefly in response to a sample writing test essay assignment and discussing the process we went through to begin answering the essay question. They then ranked actual essays and were led
through the process the panel uses to develop the rubric.

Phyllis Liston began by describing what the exam coordinators learned in the process: (1) implementing, coordinating, and gaining community-wide acceptance for exit exams is "a lot harder than it looks"; (2) communication at all levels is essential; (3) low-level mistakes can cause high-level difficulties; money when needed is found; and (4) holistic grading works well. In addition, the exam needs full administrative and faculty support. As the director of the writing competency exam, Liston found the administrative duties to be a full-time responsibility requiring personnel assistance.

Liston explained Ball State's "3/3/3" exam process. Students sign up for the exam three weeks before the exam date and are given an instruction sheet detailing the exam process, the exam question, how to prepare for the exam, and where to go to receive help preparing for the exam. On the exam date, students are given three hours to write approximately three pages in response to the exam question. Students must pass the test to graduate. After two attempts, they are required to enroll in--and repeat until they pass--an upper division writing course. The second opportunity to take the exam constitutes an automatic appeal. Exit from the course is by portfolio prepared by the students with the help of their instructors. Portfolios are evaluated by two or more readers other than the classroom teacher/coach. No student takes the exam more than twice.

John Mathew explained the training process for holistic graders by taking participants through a mini grading workshop. We read and ranked three sample essays high, middle and low. Then we read, ranked and integrated into the previous essays three more, and did the same for two additional essays. We then discussed our ranking of one of the essays in terms of its strengths and weaknesses. Finally we were presented with the six-point rubric developed by the panel for the particular exam and were asked to rate the essay.

In an actual reading, graders read ten papers at a time, assess, record and score, and pass the papers on to a second reader. Papers with scores that do not match are given to a third reader. All pass decisions are made by the University Provost under the advisement of the panel and other administrators after all exams for the quarter have been scored. The panel acknowledges a high reader calibration and suggests a main reason for it is that readers do not know the cut-off point for failing, and therefore are more objective and not sympathetically influenced to pass a border-line paper.

Linda Pelzer described the rubric design process. The panel develops a new rubric for each exam by reading and sorting all essays written for the exam into high, middle and low categories. After sorting, they discuss the categories and write about them, and they draft a six-point rubric—one that is quite detailed and descriptive and that includes specific examples from student papers to illustrate the rubric's categories. A six-point rubric is used because it eliminates a middle score and because a four-point rubric would not be specific enough to encompass the aspects of the writing they wish to assess. The panel takes care and time in designing the rubric to make it clear and specific in order for readers to reach consensus and to withstand criticism from students, parents and faculty. Rubrics are kept on file at the University library. One indicator of the success of the rubric is that students who fail the exam and wish to contest it usually reach agreement after examining the rubric and evaluating their own writing against it.

Although the writing competency examination project is bigger than the panel first anticipated, they agree that it is worth the work.

PROFICIENCY TESTING: ISSUES AND MODELS

Speakers: George Gadda, University of California, Los Angeles
Mary Fowles, Educational Testing Service, New Jersey

Introducer/Recorder: Adele Hansen, University of Minnesota

George Gadda opened the discussion with a statement concerning general issues in developing a proficiency testing program. Proficiency testing, like achievement testing, measures success in a particular domain. There are several motivations for proficiency testing: to certify individual achievement exclusive of grades, to validate a program's effectiveness, or to screen before certification of passing to the next level of instruction. The choice of purpose governs the rest of the assessment program. Proficiency tests may be used to exempt students from further work; to prove value added in a course program; to permit passage, graduation or certification; or to identify those who need further instruction.

Gadda noted that test-makers should define the domain of the test by describing the kind of written ability being assessed and that we should make a public statement concerning the criteria used for judgment. Tests used for advancement should be a well-defined part of the curriculum, with samples and grading criteria clearly described. Ideally, scorers should be those people who are testing and using the results. In addition, we need to determine what will happen to those who don't pass. Gadda noted that proficiency tests should not be a "roadblock." He concluded by stating that we should strive for high reliability and validity in our testing because proficiency tests need to withstand legal challenges.

Mary Fowles remarked that we need an increased understanding of what is to be tested and that the "community" must share the same standards. She referred to a project in Rhode Island, where a state administrator
through the process the panel uses to develop the rubric.

Phyllis Liston began by describing what the exam coordinators learned in the process: (1) implementing, coordinating and gaining community-wide acceptance for exit exams is "a lot harder than it looks"; (2) communication at all levels is essential; (3) low-level mistakes can cause high-level difficulties; money when needed is found; and (4) holistic grading works well. In addition, the exam needs full administrative and faculty support. As the director of the writing competency exam, Liston found the administrative duties to be a full-time responsibility requiring personnel assistance.

Liston explained Ball State's "3/3/3" exam process. Students sign up for the exam three weeks before the exam date and are given an instruction sheet detailing the exam process, the exam question, how to prepare for the exam, and where to go to receive help preparing for the exam. On the exam date, students are given three hours to write approximately three pages in response to the exam question. Students must pass the test to graduate. After two attempts, they are required to enroll in—and repeat until they pass—an upper division writing course. The second opportunity to take the exam constitutes an automatic appeal. Exit from the course is by portfolio prepared by the students with the help of their instructors. Portfolios are evaluated by two or more readers other than the classroom teacher/coach. No student takes the exam more than twice.

John Mathew explained the training process for holistic graders by taking participants through a mini grading workshop. We read and ranked three sample essays high, middle and low. Then we read, ranked and integrated into the previous essays three more, and did the same for two additional essays. We then discussed our ranking of one of the essays in terms of its strengths and weaknesses. Finally we were presented with the six-point rubric developed by the panel for the particular exam and were asked to rate the essay.

In an actual reading, graders read ten papers at a time, assess, record and score, and pass the papers on to a second reader. Papers with scores that do not match are given to a third reader. All pass decisions are made by the University Provost under the advisement of the panel and other administrators after all exams for the quarter have been scored. The panel acknowledges a high reader calibration and suggests a main reason for it is that readers do not know the cut-off point for failing, and therefore are more objective and not sympathetically influenced to pass a border-line paper.

Linda Pelzer described the rubric design process. The panel develops a new rubric for each exam by reading and sorting all essays written for the exam into high, middle and low categories. After sorting, they discuss the categories and write about them, and they draft a six-point rubric—one that is quite detailed and descriptive and that includes specific examples from student papers to illustrate the rubric's categories. A six-point rubric is used because it eliminates a middle score and because a four-point rubric would not be specific enough to encompass the aspects of the writing they wish to assess. The panel takes care and time in designing the rubric to make it clear and specific in order for readers to reach consensus and to withstand criticism from students, parents and faculty. Rubrics are kept on file at the University library. One indicator of the success of the rubric is that students who fail the exam and wish to contest it usually reach agreement after examining the rubric and evaluating their own writing against it.

Although the writing competency examination project is bigger than the panel first anticipated, they agree that it is worth the work.

PROFICIENCY TESTING: ISSUES AND MODELS

Speakers: George Gadda, University of California, Los Angeles
Mary Fowles, Educational Testing Service, New Jersey

Introducer/Recorder: Adele Hansen, University of Minnesota

George Gadda opened the discussion with a statement concerning general issues in developing a proficiency testing program. Proficiency testing, like achievement testing, measures success in a particular domain. There are several motivations for proficiency testing: to certify individual achievement exclusive of grades, to validate a program's effectiveness, or to screen before certification of passing to the next level of instruction. The choice of purpose governs the rest of the assessment program. Proficiency tests may be used to exempt students from further work; to prove value added in a course program; to permit passage, graduation or certification; or to identify those who need further instruction.

Gadda noted that test-makers should define the domain of the test by describing the kind of written ability being assessed and that we should make a public statement concerning the criteria used for judgment. Tests used for advancement should be a well-defined part of the curriculum, with samples and grading criteria clearly described. Ideally, scorers should be those people who are testing and using the results. In addition, we need to determine what will happen to those who don't pass. Gadda noted that proficiency tests should not be a "roadblock." He concluded by stating that we should strive for high reliability and validity in our testing because proficiency tests need to withstand legal challenges.

Mary Fowles remarked that we need an increased understanding of what is to be tested and that the "community" must share the same standards. She referred to a project in Rhode Island, where a state administrator