

Improving Assessment in University Economics

William B. Walstad

Abstract: The author discusses the following seven issues affecting assessment of undergraduates in universities: decisionmaking and the selection of tests, the use of written and oral assignments to measure learning, the characteristics of grades and portfolios for evaluating students, opportunities for self-assessment and feedback to instructors, retention of learning and the testing for higher-ordered thinking, the psychology of students in the economics classroom, and the development of new tests as public goods. The author suggests ways that economics faculty can add new dimensions to their assessment practices, improve their understanding of assessment choices, use assessment to enhance the quality of student thinking, and conduct research studies on assessment questions.

Key words: active learning, assessment, economics courses, research, testing
JEL code: A22

Assessment is an integral part of faculty teaching and student learning in university economics. Economics instructors spend a substantial amount of time evaluating student economic understanding through classroom tests, quizzes, homework, papers, and projects. They then use that information to assign course grades. Assessment, however, goes well beyond testing and grading. Instructors can use a variety of classroom assessment techniques to obtain feedback from students to identify learning problems and guide their teaching efforts (Angelo and Cross 1993). These learning activities also give students an opportunity to check their understanding and prepare themselves for class and to do assignments. Assessment is thus a multi-dimensional activity and information resource for economics teachers and students, a theme that will be emphasized throughout this article.

What follows is a discussion of seven issues that illustrate the multi-faceted nature of assessment. Economics faculty members should consider each one to understand the full range of assessment work and choices in economic education. Some of the issues, such as selecting a test or obtaining student feedback, are thought about primarily when teaching a course. Other issues, such as student portfolios or retention of knowledge, cut across courses or majors and are considered over longer time periods. Still other issues, such as student psychology

William B. Walstad is a professor of economics and director, National Center for Research in Economic Education, University of Nebraska-Lincoln (e-mail: wwalstad1@unl.edu). Helpful comments were provided by Sam Allgood, William Becker, Carol Johnston, Peter Kennedy, Ian McDonald, Ken Rebeck, Michael Salemi, and Michael Watts.

or new tests, have implications for the conduct of research in economic education. This article, however, is not designed as a practical, how-to guide to assessment in economics. Its purpose is to offer economics faculty members insights about assessment decisions at multiple levels and to indicate where opportunities exist to advance research in economic education.¹

To achieve this objective, I first introduce a framework for economic decision-making to structure the thinking about assessment choices. Many faculty members have strong opinions about how to assess student understanding in economics, so discussions of assessment practices are often quite normative and heated. The normative perspective may also influence research that uses assessment measures and make its primary goal one of promoting one teaching approach over another. A positive analysis of assessment decisions, or at least a reasoned evaluation of alternatives, often gets neglected in a crusade to promote a new approach to teaching or to student evaluation, especially if it has popular appeal.

TESTS AND DECISIONS

Certainly teaching an economics course requires an instructor to make many instructional decisions that can be analyzed from a decisionmaking perspective. Perhaps the most difficult problem is the selection of the type of classroom test that best assesses student achievement in economics. The five familiar steps in economic problem solving or decisionmaking are (1) defining the problem, (2) specifying the alternatives, (3) stating the criteria to evaluate the alternatives, (4) evaluating the alternatives, and (5) making a decision.

The specification of alternative measures of understanding can be tedious because there are so many types of classroom tests. In general, tests are either *constructed* response or *fixed* response. The former require students to supply a response to an item or series of items by writing, calculating, graphing, or some other activity (e.g., essays, short answers, and numerical or word problems). The latter ask the student to select the best answer from a fixed set of answers (e.g., multiple choice and true false).

For the type of test decision, the criteria would be a listing of what the economics instructor thinks is important in choosing a test format. Instructors usually prefer tests that are easy to construct. Economy in scoring is also a factor for some teachers, especially those with larger classes. A test type might be selected because it can include many questions to cover the test domain. Another factor to consider is the potential for bias in scoring because of the subjective nature of grading tests. Some types of tests are preferred because of the freedom of response for the student to write answers to questions or to express opinions. Opportunities for guessing or bluffing are factors that are occasionally taken into account when considering the type of test. An instructor may be concerned with how a test measures higher levels of cognitive learning.² There are other criteria, but these seven are sufficient for this illustration.

With the major criteria stated, an instructor can conduct an evaluation to see how well each alternative meets each criterion. The judgment can be made in words or with an assigned numerical score based on some type of *personal* rat-

ing (Table 1). If each criterion receives an equal weight, then it can be shown that the essay or short-answer formats are likely to have the highest total scores across all the criteria. Each criterion, however, can be weighted to reflect its relative importance to the instructor. Different weights will change the total score and the case for selecting one testing format over another. If economy in scoring were given double or triple weight, then the multiple-choice test is likely to have the highest score.

The point of this analysis is not to persuade instructors to use multiple-choice tests and abandon essays. Nor is the point to show how to assign a cardinal or ordinal ranking to a test decision. The conclusion is that there is no one best way to test. Test formats have advantages and disadvantages. A case can be made for any type of test depending on the weighting of the criteria by an instructor.³ This positive analysis can be used to explain why multiple-choice tests are used so extensively for testing and grading in principles of economics courses in the United States (Siegfried, Stiner, and Zhan 1996), probably because of their advantage in economy of scoring, coverage of the test domain, and less potential for grading bias.

WRITING AND SPEAKING

One of the most commonly used assessment methods beyond the classroom test is the use of writing assignments that allow students to express themselves in a more natural form than can be demonstrated in the restricted conditions that are imposed by essay tests. The question to be answered through writing assignments is whether students organize their knowledge and opinions into a coherent and clear statement that reflects economic thinking. Both in-class and out-of-class writing assignments provide good means for both assessing understanding and gaining insights into what students are thinking (Hansen 1998; Petr 1998).

There are many options for student writing—research papers, short essays, or problem sets. As with tests, these writing assignments can be compared in a decision matrix. Among the relevant criteria would be: (1) the range or depth of writing, (2) reinforcement of classroom or textbook discussion, (3) allowance for individual differences, (4) ease of grading, (5) opportunity for cheating or deception, (6) complaints because of time demands or difficulty, and (7) the opportunity for repeated assessment to measure student progress during a course.

Once the criteria are specified, it is possible to evaluate them against the writing alternatives (Table 2). The conclusion drawn from an equal weighting of the criteria is that problem sets have the most favorable rating across all criteria. The advantages of problem sets are that they are usually tied to textbook or course materials; they give some indication of the students' skills in writing (if they are more than quantitative); and they can be used as a repeated assessment during a course. This conclusion is consistent with the findings of Becker (1997), who reported that economics instructors in the United States often assigned problem sets. By contrast, the term paper is likely to have the least favorable rating, which may explain why it is not widely used. The major advantages of the term paper, however, are the depth of knowledge it measures and the allowance for individual

TABLE 1
Evaluating Testing Formats: Unweighted Results

Alternatives	Criteria							Total score
	Ease of test construction	Economy in scoring	Coverage of test domain	Potential for bias in scoring	Freedom of response/writing	Opportunity for guessing/bluffing	Tap higher cognitive levels	
Essay	Few items, hard to write +1	Difficult, much time -2	Narrow, few items -1	Yes -1	Yes, great +2	Yes -1	Higher, all +2	0
Short answer	Few items, easy to write +2	Relatively easy, some time -1	Narrow, few items -1	Yes -1	Yes, some +1	Yes -1	Higher, limited +1	0
Problems (verbal or numerical)	Few items, hard to write +1	Relatively easy, some time -1	Narrow, few items -1	Yes -1	No -1	No +1	Yes, limited +1	-1
Multiple choice	Many items, hard to write -2	Very easy, little time +2	Very wide, many items and flexible +1	No +1	No -1	Yes -1	Varies, mostly lower -1	-1
True-false	Many items, easy to write -1	Very easy, little time +2	Very wide, many items and flexible +1	No +1	No -1	Yes, great -2	Lower -2	-2

VIP

TABLE 2
Evaluating Writing Assignments: Unweighted Results

Alternative	Criteria							Total score
	Range or depth of writing	Ease of grading	Direct reinforcement of course material	Allows for individual choice or differences	Potential for problems or student reactions	Opportunity for cheating	Capable of repeated assessment	
Research or term paper	Wide range or depth +2	Difficult, much time -2	Limited, focus on special topic -2	Yes, great. Select own topic. +2	Yes, great. Help select, write, revise. -2	Some, buy papers, plagiarize. -1	Little. Done once a course. -1	-4
Common short paper or essay	Limited range, some depth +1	Easier, but some time -1	Yes, some, if topic is related +1	Not much. All do same topic. -1	Yes, some, if there are writing problems. -1	Yes, great. All do same work. -2	Yes, some. 1-2 times a course. +1	-2
Problem sets or written exercises	Limited range, but some writing +1	Easier, but some time -1	Yes, great. Tied to course material. +2	No -2	No, students are familiar +1	Yes, great. all do same work. -2	Yes, great. many times during a course. +2	+1

choice. If these criteria were given more weight, then term papers would be assigned more frequently. The rating on these criteria probably accounts for why term papers are used with high ability students and in smaller classes where faculty members have more time to direct and evaluate the student work.

Being able to speak intelligently about economics is as important as being able to write about it. In fact, speaking ability may be more useful for students because they are more likely to have to speak about economic issues than write about them. Many assignments can be used to improve the oral expression of economic ideas. Case studies or readings can be given to students to discuss in groups and to make presentations in class (Hansen and Salemi 1998; Buckles 1998). Simulations can be conducted as active learning activities and to give feedback on economic understanding (Holt and McDaniel 1998). Group projects can involve the preparation of reports for oral presentations (Bartlett 1998; Cameron 1998).

The opportunities for oral communication to develop economic understanding are as unlimited as those for using writing. The major problem that weighs against the use of oral assignments and activities is that it is costly for the instructor to evaluate. It takes a substantial amount of time to listen to each student or group report and develop a valid and reliable rubric for assessing students. The time cost and uncertainty about what to evaluate probably explain why oral presentation of economic ideas is often not assessed in courses.⁴

The need for writing and speaking to show mastery of economics has long been recognized as being important for economics learning. Hansen (1986, updated in Hansen 2001) outlined a thoughtful set of proficiencies for the economics major that move from lower to higher levels of cognitive thinking. These proficiencies require use of writing and speaking assignments to show that they have been met. The drawback is that course work in economics does not provide good opportunities for developing writing and speaking skills in economics. Siegfried (1998, 67) concluded from a review of the economics major in the United States that “the amount and type of student writing assignments and oral presentations in many programs not only fail to prepare students for the demands they will encounter after graduation, but they also limit the ability of students to demonstrate their mastery of economics while still in college.” This strong criticism should be addressed through innovative instruction and assessment.

GRADES AND PORTFOLIOS

A major purpose of administering classroom tests and course assignments is to obtain information on students that can be used to assign grades. Grading is a process that gives different weights to different aspects of student work, so grades are essentially a composite measure composed of multiple indicators of student performance. Tests are one important indicator for assigning grades. Tests can even be considered multiple indicators when they contain a combination of fixed- and constructed-response items. Similarly, course assignments can vary widely in format and difficulty and can thus provide multiple indicators of student work. Other factors, such as student effort or participation, can be taken into account in grading.

The use of multiple indicators for assigning grades has several attractive attributes. It allows instructors to accommodate, at least in part, the preferences students might have for taking a type of test or for being graded on work done outside of class rather than just on course exams taken in class. It also permits instructors to address concerns about error in measurement occurring from reliance on one indicator to make a grade judgment about student achievement or performance in an economics course. Perhaps these reasons underlie the decision of many faculty members in the United States to collect several types of information on students for use in determining grades, according to survey studies (Siegfried et al. 1996).

Nevertheless, economics instructors and researchers in economic education need to give more attention to how grades are determined. Grading schemes that use different weights for test and course assignments may produce different rankings of students. How instructors determine the weighting of grade components deserves further study to evaluate the match with instructional objectives for the course. If an instructor wants to be sure that students know many concepts and can apply the concepts to new situations, then fixed-response tests can handle that objective. If instructors want students to express economic ideas through writing within time limits, then essays or short-answer questions are useful. If instructors want students to write and speak about a range of economic topics, then papers and presentations will be beneficial. What must be better understood are the instructional objectives that are being targeted in a course and the multiple indicators, weights, and grades that are being used to determine whether those instructional objectives are being met.

Assessing the economics major is a more difficult challenge than assessing student performance in a particular economics course. In most cases in the United States, students are certified as having majored in economics if they complete a prescribed set of economics courses with passing grades. The question to be asked, however, is whether there are qualitative attributes or proficiencies that are associated with majoring in economics that should be demonstrated by students prior to being awarded an economics major.

Assessment using student portfolios offers a solution for evaluating whether a student has the attributes or proficiencies of an economics major. Portfolio assessment involves evaluating a representative collection of student work that provides examples of both progress-in-learning and achievement (Arter and Spandel 1992). The evidence of progress-in-learning could be in the form of a reflective journal, or beginning-of-course and end-of-course essays on an economic issue. The achievement evidence can include a collection of tests, papers, videotape presentations, or other assignments that the student selects from different courses.

Portfolio assessment offers a way to address criticisms of writing and speaking assignments in economics courses that were previously cited and the recommendation of Hansen (1986; 2001) to measure student proficiencies through a series of writing and speaking assignments. Economics departments can establish their own set of proficiencies that require writing, speaking, or other skills (e.g., data analysis). The specification of the proficiencies can encourage faculty

Not course

to include a variety of assignments in their courses that help the student meet the proficiencies. Students can also be given a suggested list of items to include in the portfolios that are drawn from their economics courses that show that they have met the department's standards for proficiencies. Economics instructors can evaluate the portfolios using a pass-fail or letter grade system. Those students with deficiencies would be given time to prepare new items or revise previous work. The promise of portfolio assessment, especially for evaluating economics majors, merits further exploration.

SELF-ASSESSMENT AND FEEDBACK

At least three major purposes for assessment in economics courses should be carefully considered by instructors. So far the focus has been on the use of assessment for measuring student achievement and/or assigning grades because it is usually required of all instructors. Two other purposes, however, deserve equal attention if there is to be improvement in student learning and faculty teaching in economics. Instructors need to design in-class and out-of-class projects that give students frequent opportunities for self-assessment.⁵ At the same time, these projects will give feedback to instructors that can be used to increase the effectiveness of their teaching.

Many techniques for classroom assessment can be used in economics courses to meet these two purposes. Students' opinions or knowledge about an economic issue can be obtained both before and after the issue is discussed in class. Sample questions from a test can be used for class discussion or group work to prepare students for a course exam. Students can be asked to summarize in a written paragraph the main point of a lecture, newspaper article, or video clip that can then be checked by other students or the instructor for clarity and content accuracy. Oral explanations can be requested from students of the meaning of a table, graph, headline, or cartoon related to economics.⁶

As a longer assignment for self-assessment, the instructor can ask students to keep a journal about current events in economics. It would include news articles on economic issues that would be annotated to explain the economic concepts and give students the opportunity to provide their opinions about the issues. The journals serve as a form of dialogue on economics between the students and instructor and as a source of ideas for short papers or essays. The instructor can grade journal entries if content is a major component, selectively grade, or not grade if the purpose is to obtain insights into student opinions (Petr 1998).

The concern that faculty members often express about self-assessment activities is how seriously students will take them. The incentive problem can be easily overcome if the teacher explains to the students at the outset how these projects help them check their understanding of the course material and get them to focus on essential or difficult ideas. Students also need to see how these activities are useful for preparing them for tests and assignments that will be graded. If the incentive problem is thought to be severe, it is also possible to use some form of selective grading. Miller and Westmoreland (1998) selectively graded only one or two problems in homework assignments and found this combined

scheme of graded and self-assessment had no adverse effects on the number of assignments submitted or their quality.

Actions taken in a course to improve the self-understanding of students also benefit teachers by giving them insights about what students are thinking. These insights can be used to improve teaching effectiveness, student achievement, and attitudes toward economics instruction. Research on this feedback topic in economic education is still in the initial stages. In a pilot study, principles of economics students in an experimental group wrote a “one-minute paper” at the end of each class to state one important thing they learned and one item that most confused them (Chizmar and Ostrosky 1998). The findings showed the use of the one-minute technique made a contribution to economics achievement in this principles course.

RETENTION AND HIGHER THINKING

A puzzling result from research is that the economics instruction that students receive at the university level seems to have little effect on what they know about basic economics when they graduate from a university or afterward. An early critic of instruction in principles courses was Stigler (1963), who posited that if an essay test on current economic problems was given to graduates five years after attending a university, then there would be no difference in test performance between alumni who had taken a “conventional” one-year course in economics and those who had never taken a course in economics.

The Stigler hypothesis has never been tested with essays, but it has been evaluated with multiple-choice questions. In one study of the lasting effects of economic instruction, a 33-item test was administered to samples of university alumni who had and who had not taken a two-semester course in introductory economics when they were sophomores (seven years previously). The alumni who had taken these introductory courses scored 3.2 points higher than the alumni who had not taken these courses, after controlling for other factors. University seniors who had taken these economics courses as sophomores (two years previously) scored 4.8 points higher than university seniors who had not taken economics. A recent study of college seniors found similar results, with a 2-point difference on a 15-item test in the scores of seniors who had or had not taken an economics course (Walstad and Allgood 1999).

These findings can be interpreted in two ways. First, the statistically significant difference in economics test scores favoring alumni or college seniors who took a university economics course casts doubt on the Stigler hypothesis. Second, the test score difference is minimal, even if it is statistically significant, and the final level of achievement is low. In this case, the hypothesis and its implied criticism of principles instruction cannot be dismissed.

If the second interpretation is accepted, then there are implications for assessment. Testing for understanding of basic economic concepts and relationships may be sufficient when evaluating student learning in most courses and programs (not only in studies of lasting effects). If students cannot answer questions about basic concepts in economics, then it is inefficient to test economic thinking at

higher cognitive levels. For example, if students do not know the difference between a movement along the demand curve and a change in demand, then asking them in an essay question to explain an economic event using supply and demand analysis is likely to produce a confused or incomplete answer. If, however, students do show mastery of basic economics, then testing for higher-ordered thinking should be done.

This progression is also suggested by the low-to-high taxonomies of Bloom et al. (1956) and Hansen (1986; 2001). Students first need to show a certain degree of mastery of basic economic concepts because that knowledge and understanding lay the foundation for higher-ordered thinking about economics. A similar statement can be made about assessment in mathematics or the physical sciences. Helping students to “think like an economist” (Siegfried 1991, 199) is often considered to be the goal of economics instruction at the university level, but developing that thinking implies an understanding of basic economic concepts (Siegfried 1998).

PSYCHOLOGY AND ECONOMIC EDUCATION

Student thinking or perceptions about economics and their courses deserve further investigation by teachers and researchers in economic education. One study in a principles of economics course found that the overall grade point average that students self-reported was often higher than those that were recorded in university records (Maxwell and Lopus 1994). Other studies have found that self-reports of study time are often unrelated to student achievement (Becker 1982). Students may also be entering courses with grade expectations that are too optimistic relative to the distribution of course grades (Falchikov and Boud 1989).

The deeper issue is that there may be nonrational behavior and thinking that are affecting student achievement or course completion. Perhaps it is time to start explaining the psychology of students in economics courses using some of the models and findings from behavioral economics (Rabin 1998) that have been applied to explain the psychology of investors in the stock market. The logical leap may not be great because students are more likely to view economics courses as investment rather than consumption. Concepts from behavioral economics, such as overconfidence, the law of small numbers, confirmatory biases, random anchoring, asymmetric loss aversion, and time-variant preferences (Rabin 1998) may be useful for explaining some of the anomalies in student behavior in classes, such as low attendance, inadequate preparation, poor test performance, and student complaints. These insights could be used to prevent misperceptions or misunderstandings before they develop in the minds of students. The application of this research to questions about the thinking of students in economics may be relatively easy because most of the research in behavioral economics is conducted with experiments using different groups of people (including students). Experiments are not new to economic education (Holt and McDaniel 1998).

Assumptions can also be made that students act rationally. Some traditional models have been used to explain anomalies, such as the lack of the effect of study time on achievement. Becker (1982) specified a model of student utility

maximization and found that changes in instruction that permit students to convert time into learning more efficiently need not result in any change in the students' learning of economics. The behavioral response is unpredictable because the benefits from more efficient instruction that reduce study time may be used by students to increase learning in other subjects or may be converted into more leisure. Similarly, Allgood (forthcoming) specified a utility maximizing model that assumes that students target grades. Once a student achieves an acceptable grade threshold, then the student's effort in the course can wane. These economic models can be used to explain why instructional innovations are often not effective, why attendance is poor even among high-ability students, or why the degree of student effort varies during a course.

NEW TESTS AND PUBLIC GOODS

Since 1975, one standardized test has been consistently used for research in economic education at the university level—the Test of Understanding of College Economics (TUCE) (Saunders 1991). In its three editions, it has served as the outcome measure for several hundred studies. The test, however, is not without controversy because it has been criticized for its content coverage and the fact that it is a multiple-choice test.

The interesting question is why other standardized measures of student achievement in economics at the university level, either in fixed-response or constructed-response formats, have not been created for research in economic education.⁷ It is not as if the problem has not been recognized because over the years several researchers have called for the preparation of new tests of economics learning. Although other measures, such as persistence in taking economics courses and attrition from courses, have been used to assess outcomes from instruction (Becker 1997), no new standardized tests have been developed for research.

A likely explanation for the lack of test development in economics is that the product of the work has some of the characteristics of a public good. There is not much incentive to construct test measures with good documentation on reliability and validity given the limited market demand for such a product. It is more exciting and rewarding (from the perspective of a researcher or faculty member) to conduct studies using a test than it is to develop one. Once such a measure is available for assessment, however, the test users can obtain the benefits without paying for the development costs. The situation makes the creation of quality measures almost impossible unless there is grant funding to underwrite the development costs.

It is difficult to conduct comparative research studies of economic achievement without valid and reliable tests. The problem with classroom tests is that little information is available for judging measurement properties. The same problem applies to grades, although it is often ignored in practice. Studies that aggregate grades across courses, instructors, and institutions to measure achievement create measurement problems because of different grading scales. An "A" grade in one economics course may not have the same meaning as an "A" grade in another

course. Standardized measures of achievement are needed that overcome these limitations so better studies can be conducted within and across institutions.

Alternatives to multiple-choice tests such as the TUCE are also needed if research in economic education is to be advanced. Performance assessment is often considered to be a better alternative and a more direct way to measure important outcomes from instruction.⁸ It generally refers to tasks that require students to show their understanding or mastery of a subject through some kind of written or oral project. These performances can include taking an essay test, writing a short or long paper, preparing a written or multimedia report, participating in a debate, or making an oral presentation. The purpose of this assessment is for students to produce a product to demonstrate that the goals for economic instruction have been met.

The major problem with performance assessment is that it must still meet the basic standards of reliability and validity. Journals in educational measurement have devoted much space to articles examining the technical qualities of alternative measures. Brennan and Johnson (1995, 12) highlight the measurement dilemma when they state that “the realism of performance assessments comes at the cost of limitations in the generalizability of the results.” The lesson for economic education from this work is that the adoption of new labels for tests will not result in better assessment of economics students. Calling a new outcome measure in economics a performance, authentic, critical thinking, or higher-ordered test does not mean that it has those desired features. Substantial work is required to produce a standardized measure of performance outcomes, especially if it is to be used for comparative testing across classes and institutions and for research studies.⁹ This caution is not meant to imply that alternative measures should not be developed, but rather that more resources need to be devoted to creating these new tests and documenting their use with students. New standardized tests in economics will benefit from the past studies in other subjects.

CONCLUSION

Several conclusions can be drawn from this broad discussion of assessment in university economics. First, assessment needs to be viewed as multi-dimensional. Many forms of assessment—for testing and grading, as feedback to the instructor and for self-assessment by students—can and should be used by economic instructors in their courses to improve student learning. Second, decisions on the best way to conduct assessment with students depend on the instructor’s evaluation of the relative merits of each alternative in light of the criteria. Each assessment decision should be viewed from a positive rather than a normative perspective to help the instructor think more deeply about the choice to be made and justify the outcome. Third, the quality of the assessment of students in economics courses is a reflection of the quality of teaching in these courses. By asking students to speak, to write, or to take a test on economics, the instructor is asking the students to think about economics. The higher the quality of the assessment (both graded and ungraded), the more likely it will be that students will do more thinking in a course and increase their learning. Fourth, assessment

in economics can be improved with more research studies, both large and small. More understanding is needed of what students retain over time from taking an economics course. More attention should be given to student behavior in economics courses and how assessment practices affect it. New standardized tests in economics should be developed to measure outcomes from economics courses and for evaluations of teaching innovations in economics.

NOTES

1. For related discussion of research in teaching undergraduate economics, see Becker (1997).
2. Bloom et al. (1956) developed one such taxonomy with six levels: knowledge, comprehension, application, analysis, synthesis, and evaluation.
3. Other factors or qualities may influence the type of test chosen and the ratings. Multiple-choice tests may be easy to write if test banks are used, and it is possible to write multiple-choice items to assess understanding at higher cognitive levels. An advantage of numerical problems is that there is less opportunity for guessing, so an instructor concerned about problem-solving mastery may give more weight to this criterion. Most testing types can be modified to address important criteria. For example, an instructor can overcome the disadvantage with guessing on true-false items if students are asked to give written rationales as answers.
4. A decision table similar to the one in Table 1 could be constructed for the speaking decision, but for the sake of parsimony, it will not be presented. The decision table is a useful device for highlighting the normative and positive perspectives of any assessment decision.
5. Crooks (1988) argues that too much emphasis has been devoted to the grading function of assessment. He thinks that students should have frequent opportunities to check for self-understanding to show mastery of material at a deeper level, and that students should be given high but attainable standards.
6. Buckles (1998), Cameron (1998), Hansen (1998), Petr (1998), and Walstad and Saunders (1998) provide examples of active learning techniques for self-assessment and feedback in economics. Angelo and Cross (1993) offer a general handbook on these techniques.
7. In the United States, tests such as the Advanced Placement (AP) exam in economics, the Graduate Record Exam (GRE) in economics, and the Major Field Test (MFT) in economics were developed for commercial purposes. The GRE and the MFT are multiple-choice tests. The AP exam is composed of two-thirds multiple choice and one-third essay or short answer. These commercial tests cannot be easily used for research studies that require flexibility.
8. It is common practice in education to attach adjectives to assessment. Caution, however, should be exercised in using these terms because they are vague or there can be an agenda attached. For example, Terwilliger (1997) comments that the use of "authentic," or labels such as "genuine" or "realistic" are misleading and confusing because they are based on a value judgment and imply that any other alternative is not authentic, genuine, or realistic. He concludes that claims for authentic assessment are "based largely on an appeal to face validity, a concept that has been abandoned by modern psychometric theorists" (p. 27).
9. Miller and Legg (1993) discuss measurement challenges with alternative assessment including validity, reliability, equating, fairness, and the effects on both teaching and learning.

REFERENCES

- Allgood, S. forthcoming. Grade targets and teaching innovations. *Economics of Education Review*.
- Angelo, T. A., and K. P. Cross. 1993. *Classroom assessment techniques: A handbook for college teachers*, 2nd ed. San Francisco: Jossey-Bass.
- Arter, J. A., and V. Spandel. 1992. Using portfolios of student work in instruction and assessment. *Educational Measurement: Issues and Practice* 11 (1): 36–44.
- Bartlett, R. L. 1998. Making cooperative learning work in economics classes. In W. E. Becker and M. Watts, eds., *Teaching economics to undergraduates: Alternatives to chalk and talk*, 11–34. Cheltenham, U.K.: Edward Elgar.
- Becker, W. E. 1982. The educational process and student achievement given uncertainty in measurement. *American Economic Review* 72 (1): 229–36.
- . 1997. Teaching economics to undergraduates. *Journal of Economic Literature* 35 (3): 1347–73.

- Bloom, B. S., M. D. Englehart, E. J. Furst, W. H. Hill, and D. R. Krathwohl. 1956. *Taxonomy of educational objectives handbook I: Cognitive domain*. New York: David MacKay.
- Brennan, R. L., and E. G. Johnson. 1995. Generalizability of performance assessments. *Educational Measurement: Issues and Practice* 14 (4): 5–17.
- Buckles, S. 1998. Using cases as an effective active learning technique. In W. E. Becker and M. Watts, eds., *Teaching economics to undergraduates: Alternatives to chalk and talk*, 225–40. Cheltenham, U.K.: Edward Elgar.
- Cameron, B. 1998. Active and cooperative learning strategies for economics classrooms. In W. B. Walstad and P. Saunders, eds., *Teaching undergraduate economics: A handbook for instructors*, 245–56. New York: Irwin/McGraw-Hill.
- Chizmar, J. F., and A. L. Ostrosky. 1998. The one-minute paper: Some empirical findings. *Journal of Economic Education* 29 (Winter): 3–10.
- Crooks, T. J. 1988. The impact of classroom evaluation practices on students. *Review of Educational Research* 58 (4): 438–89.
- Falchikov, N., and D. Boud. 1989. Student self-assessment in higher education: A meta-analysis. *Review of Educational Research* 59 (4): 395–430.
- Hansen, W. L. 1986. What knowledge is most worth knowing—For economics majors? *American Economic Review* 76 (2): 149–52.
- . 1998. Integrating the practice of writing into economics instruction. In W. E. Becker and M. Watts, eds., *Teaching economics to undergraduates: Alternatives to chalk and talk*, 79–118. Cheltenham, U.K.: Edward Elgar.
- . 2001. Expected proficiencies for undergraduate economics majors. *Journal of Economic Education* 32 (3): 231–42.
- Hansen, W. L., and M. K. Salemi. 1998. Improving classroom discussion in economics courses. In W. B. Walstad and P. Saunders, eds., *Teaching undergraduate economics: A handbook for instructors*, 207–26. New York: Irwin/McGraw-Hill.
- Holt, C. A., and T. McDaniel. 1998. Experimental economics in the classroom. In W. B. Walstad and P. Saunders, eds., *Teaching undergraduate economics: A handbook for instructors*, 257–68. New York: Irwin/McGraw-Hill.
- Maxwell, N. L., and J. S. Lopus. 1994. The Lake Wobegon effect in student self-reported data. *American Economic Review* 84 (2): 201–05.
- Miller, E., and G. Westmoreland. 1998. Student response to selective grading in college economics courses. *Journal of Economic Education* 29 (3): 195–201.
- Miller, M. D., and S. M. Legg. 1993. Alternative assessment in a high-stakes environment. *Educational Measurement: Issues and Practice* 12 (2): 9–15.
- Petr, J. 1998. Student writing as a guide to student thinking. In W. B. Walstad and P. Saunders, eds., *Teaching undergraduate economics: A handbook for instructors*, 227–43. New York: Irwin/McGraw-Hill.
- Rabin, M. 1998. Psychology and economics. *Journal of Economic Literature* 36 (1): 11–46.
- Saunders, P. 1980. The lasting effects of introductory economics courses. *Journal of Economic Education* 12 (1): 1–14.
- . 1991. *Test of understanding of college economics: Examiner's manual*, 3rd ed. New York: National Council on Economic Education.
- Siegfried, J. J. 1998. The goals and objectives for the economics major. In W. B. Walstad and P. Saunders, eds., *Teaching undergraduate economics: A handbook for instructors*, 59–72. New York: Irwin/McGraw-Hill.
- Siegfried, J. J., P. Saunders, E. Stinar, and H. Zhang. 1996. How is introductory economics taught in America? *Economic Inquiry* 34 (1): 1182–92.
- Stigler, G. 1963. Elementary economic education. *American Economic Review* 53 (2): 653–59.
- Terwilliger, J. 1997. Semantics, psychometrics, and assessment reform: A close look at “authentic” assessment. *Educational Researcher* 26 (8): 24–27.
- Walstad, W. B., and S. Allgood. 1999. What do college seniors know about economics? *American Economic Review* 89 (2): 350–354.
- Walstad, W. B., and P. Saunders. 1998. Using student and faculty evaluations to improve economics instruction. In W. B. Walstad and P. Saunders, eds., *Teaching undergraduate economics: A handbook for instructors*, 337–53. New York: Irwin/McGraw-Hill.