International Statistical Review (1999), 67, 1, 1-12, Printed in Mexico

Assessment and Statistics Education: Current Challenges and Directions

Joan B. Garfield[1] and Iddo Gal[2]

[1]*Department of Educational Psychology, 332 Burton Hall, University of Minnesota, USA*
[2]*Department of Human Services, 512 Eshkol Tower, University of Haifa, Israel*

Summary

The interaction between new curricular goals for students and alternative methods of assessing student learning is described. Suggestions are offered for teachers of statistics who wish to re-examine their classroom assessment practices in light of these changes. Examples are offered of some innovative assessment approaches that have been used in introductory statistics courses, and current challenges to statistics educators are described.

Key words: Assessment; Reasoning; Learning.

1 Introduction: The Assessment Revolution

Several years ago an "assessment revolution" began to transform the education world, affecting all subject areas. Dissatisfaction with traditional approaches to educational assessment was a leading force behind this movement. A general consensus developed that some typical forms of assessment (e.g., multiple-choice exams) are too narrow to provide sufficient information about student learning. Traditional forms of assessment began to be viewed as not aligned with current instructional and curricular goals, and single number scores or letter grades assigned to students were realized to be an inadequate way to evaluate student understanding or to promote successful learning.

As part of the "assessment revolution", the definition of "assessment" has changed. Traditionally, people think of the words "test" and "grade" when they hear the term "assessment". Now, a shift has taken place from an emphasis on "testing" to an emphasis on "assessment" and to the integration of assessment and instruction (Birenbaum & Dochy, 1996). Newer definitions of assessment refer to procedures more likely to elicit complex intellectual performance than traditional paper and pencil tests. Some of these alternative procedures now include portfolios of students' work, student projects, and laboratory activities (Darling–Hammond *et al.*, 1995).

In the past, assessment in statistics courses consisted primarily of items measuring mastery of skills, procedures, and vocabulary, or items that test skills in isolation of a problem context. These types of items are rarely able to test whether or not students understand statistical concepts, are able to integrate statistical knowledge to solve a problem, or are able to communicate effectively using the language of statistics. This approach to assessment is now changing, as more students are assessed in their courses using alternative methods, as described in this paper.

2 New Goals for Statistics Learners

As part of an educational reform movement that is affecting the statistics community, new ideas have emerged about how students should learn statistics, what topics are most important to learn, and how these topics should be taught (Moore, 1997). New goals have emerged for students enrolled in statistics courses that place an emphasis on analyzing data, interpreting data, and communicating about data and chance. An examination of the curricular goals reveals inadequacies inherent in exclusively using traditional forms of assessment to measure student learning.

Gal & Garfield (1997) provide a summary of currently accepted learning goals for students in statistics courses, which include:

1. Understand the Purpose and Logic of Statistical Investigations

Students should understand why statistical investigations are conducted, and the "big ideas" that underlie approaches to data-based inquiries. These ideas include:

- The existence of variation.
- The need to describe populations by collecting data.
- The need to reduce raw data by noting trends and main features through summaries and displays of the data.
- The need to study samples instead of populations and to infer from samples to populations.
- The logic behind related sampling processes.
- The notion of error in measurement and inference, and the need to find ways to estimate and control errors.
- The logic behind methods (such as experiments) for determining causal processes.

2. Understand the Process of Statistical Investigations

Students should understand the nature of and processes involved in a statistical investigation and considerations affecting the design of a plan for data collection. They should recognize how, when, and why existing statistical tools can be used to help an investigative process. They should be familiar with the specific phases of a statistical inquiry. These phases include:

- Formulating a question.
- Planning the study (e.g., approach and overall design, sampling, choice of measurement tools).
- Collecting and organizing data.
- Displaying, exploring, and analyzing data.
- Interpreting findings in light of the research questions.
- Discussing conclusions and implications from the findings, and identifying issues for further study.

3. Master Important Procedural Skills

Students need to master the "component skills" that may be used in the process of a statistical investigation. This mastery includes being able to organize data, compute needed indices (e.g., median, average, confidence interval), or construct and display useful tables, graphs, plots, and charts, either by hand or assisted by technology (e.g., a calculator, graphing calculator, or computer).

4. Understand Probability and Chance

Moore (1997) recommends that students need only an informal grasp of probability in order to follow the reasoning of statistical inference. This understanding would develop from experiences with chance behavior starting with devices (e.g., coins and dice) and leading to computer simulations. In this way students should gain an understanding of a few key ideas. These ideas include:

- Concepts and words related to chance, uncertainty, and probability that appear in our everyday lives, particularly in the media.
- It is important to understand probabilistic processes in order to better understand the likelihood of events in the world around us, as well as information in the media.
- Probability is a measure of uncertainty.
- Developing a model and using it to simulate events is a helpful way to generate data to estimate probabilities.
- Sometimes our intuition is incorrect and can lead us to the wrong conclusion regarding probability and chance events.

5. Develop Interpretive Skills and Statistical Literacy

In carrying out a statistical investigation, students need to be able to interpret results and be aware of possible biases or limitations on the generalizations that can be drawn from data. We realize that most students are more likely to be consumers of data than researchers, and will seldom have to collect or analyze data as adults. Instead, they will need to be able to make sense of published results from studies and surveys reported in the media or in a workplace context. Therefore, students need to learn what is involved in interpreting results from a statistical investigation and to pose critical and reflective questions about arguments that refer to summary statistics or to data reported in the media or in project reports from their classroom peers (e.g., How reliable are the measurements used? How representative was the sample? Are the claims being made sensible in light of the data and sample?).

6. Develop Ability to Communicate Statistically

Strong writing and speaking skills are needed if students are to effectively communicate about statistical investigations and probabilistic phenomena or processes. Good reading comprehension and communication skills are required so that students can effectively discuss or critique statistical or probabilistic arguments they encounter which claim to be based on some data (e.g., "8 out of 10 doctors use . . . ", "there is a 20 percent chance that . . . "). Students should be able to use statistical and probabilistic terminology properly, convey results in a convincing way, and be able to construct proper arguments based on data or observations. They should also be able to argue thoughtfully about the validity of other people's interpretations of data or graphical displays, and raise questions about acceptability of generalizations made on the basis of a single study or a small sample.

7. Develop Useful Statistical Dispositions

Students should develop an appreciation for the role of chance and randomness in the world and for statistical methods and planned experiments as useful scientific tools and as powerful means for making personal, social, and business-related decisions in the face of uncertainty. They should realize that the process of statistical inquiry can often lead to better conclusions than relying on anecdotal data or on their own subjective experiences or intuitions, but that this is not guaranteed. Students should also learn to adopt a questioning stance when they are faced with an argument that purports to be based on data (e.g., "all people are . . . ") or a report of results or conclusions from a

statistical investigation, survey, or empirical research.

3 The Need for Alternative Approaches to Assessment

The expanded goals for student learning have led to the exploration and development of alternative types of assessment, for use in classes, research, and curriculum evaluation. These efforts have been built on principles that assessment should reflect the statistical content that is most important for students to learn and that assessment should enhance learning of statistics and support good instructional practice.

The use of the new, alternative forms of assessment appears to have many advantages. For example, alternative assessments often provide more complete information about what students have learned and are able to do with their knowledge, and tend to focus more on processes than on products. They may provide more detailed and timely feedback to students about the quality of their learning and better capture how students think, reason, and apply their learning (rather than merely having students "tell" the teacher what they have remembered or show that they can perform calculations or carry out procedures correctly). Use of multiple methods rather than a single test also provides a richer and more complete representation of student learning.

4 Innovative Models for Classroom Assessment

What do some of the alternative forms of assessment look like? Many of them now are quite familiar, and may seem typical, rather than "alternative". They include:

1. Individual or Group Projects

Projects completed by individual students or groups of students typically involve posing a problem, designing an experiment or taking a sample, collecting and analyzing data, and interpreting the results. The project may be written up as a report, presented orally in class, or displayed on a poster. Projects may be assessed using a scoring rubric to assign points (such as 0, 1, 2) to different components of the project. For example, the following categories may be used to evaluate a student project and scores of 0 through 2 would be assigned to each category:

1. Demonstrates understanding of the problem being addressed.
2. Uses appropriate methods to collect the data.
3. Uses appropriate methods to analyze the data.
4. Provides an adequate interpretation of the data analysis.
5. Discusses limitations of the project.
6. Communicates effectively in the written report.

In this example, a maximum score of 12 would indicate the highest level of achievement.

For further information on projects see Mackisack (1994) or Starkings (1997).

2. Portfolios of Student Work

Portfolios consist of a collection of students' work, often gathered over an entire course. The selection is often done by both the student and teacher and may include a variety of components, such as computer output for data analyses, written interpretations of statistical analyses, and reflections on what has been learned. Keeler (1997) describes the use of portfolios in a statistics class for graduate students. In this example, portfolios include:

1. A mini research paper, where students investigate a problem or question that can be answered by collecting binomial data.
2. Six SAS and SPSSX programs the students have used and the output for these programs.
3. A group project involving data collection and analysis, that has been presented to the class.
4. Two midterm exams accompanied by students' analysis of their own performance and their corrections of wrong answers.
5. A reflective journal.

As each assessment task is entered into the appropriate section of the portfolio, students are asked to respond to the following questions:

- What did you have trouble with in producing this product?
- How did you change the product after the initial feedback?
- How could the product still be improved? and
- What was the most important thing you learned in producing this product?

The portfolio is awarded 300 total points which is about one-third of the total points used in computing a course grade. Points are awarded to each section of the portfolio according to a detailed scoring rubric (see Keeler, 1997).

3. Concept Maps

Schau & Mattern (1997) describe different uses of concept maps to assess students' understanding of conceptual connections. Concept maps include the concepts (referred to as nodes and often represented visually by ovals or rectangles) and the connections (referred to as links and often represented with arrows) that relate them. Students may be asked to construct their own maps for a particular statistical topic (e.g., hypothesis testing) or to fill in missing components from a partially constructed concept map. The general process for this second approach involves first constructing a master map. Keeping that map structure intact, some or all of the concept and/or relationship words are omitted. Students fill in these blanks either by generating the words or by selecting them from a list which may or may not include distracters. For examples of this type of assessment see Schau & Mattern (1997).

4. Critiques of Statistical Ideas or Issues in the News

Students may be asked to read and critique a newspaper article responding to particular questions as shown in Figure 1. Points may be awarded based on the completeness of each answer.

For examples of assessment items that involve brief articles or graphs presented in the media see Watson, 1997.

5. Objective Format Questions to Assess Higher Level Thinking

Enhanced multiple-choice items or items that require students to match concepts or questions with appropriate explanations, may be used to capture students' reasoning and measure conceptual understanding. Cobb (1998) offers five principles for designing objective format questions that assess statistical thinking. One principle is to ask for comparative judgments, not just category matching. For example, a set of two-way tables is presented to students with data representing factors related to the death sentence. Each table displays frequencies for the breakdown of different independent variables (e.g., race of defendant, race of victim, prior record) by the same dependent variable (whether or not a convicted murderer is sentenced to death). Students are asked which factors in the tables are most strongly associated with whether a convicted murderer is sentenced to death.

Review of Article in the News

Select an article in a recent newspaper or magazine that particularly interests you and that reports on the results of some type of research study. Be sure that the article you select provides enough information so that you are able to answer the following questions (which should be used as headings for your report). Attach a copy of the article.

Write-up:

1. What do you think is the purpose of the research study described in this article?
2. What method or methods were used to answer the research question?
3. What questions would you like to ask the investigators in order to better understand the study?
4. Are there any aspects of the study that might make you question the conclusions presented in the article?

Figure 1. *An assessment activity involving critiquing statistical ideas in the news*

They need to compare the strength of the interaction between variables in each table to make this judgment. A second principle is to involve two or more modes of statistical thinking (e.g., both visual and verbal/intuitive thinking). An example is given where students are asked to match verbal descriptions to four different plots of data.
For further information on objective-format tests, see Hubbard (1997), Jolliffe (1997), or (Wild, Triggs & Pfannkuch, 1997).

6. Minute Papers

Minute papers are brief, anonymous written remarks provided by students, sometimes on an index card or half-sheet of paper, during the last few minutes of class. These remarks can cover a variety of topics, such as a summary of what students understand or do not understand on a topic, or students' reactions to various aspects of a course (e.g., the use of cooperative groups, the textbook, or the teacher's explanations in class). Some statistics teachers use minute papers to have students to describe their understanding of a particular concept or procedure discussed in class that day, or to have them respond to the question: "what was the most confusing idea in today's class?" Another form of a minute paper is an open-ended student evaluation of the course, while the class is in progress. An example is shown in figure 2.
Students' comments on these surveys may be tallied and sorted, and a list of things that can be responded to and that are possible (and reasonable) to change may emerge. The feedback from students may sometimes lift a teacher's spirits, and often maintains a teacher's humility. The ultimate goal of obtaining this information on students' learning, attitudes, and understanding, is to build a better understanding of the students' perceptions and experiences in the course and to suggest ways to improve teaching, making it a worthwhile endeavor despite the possible discomfort. For further information on minute papers, see Angelo & Cross (1993).

Midterm feedback

What are your reactions to this course so far?

How well do you feel you are doing?

What do you like most?

What do you like least?

Please add any suggestions or comments you have.

Figure 2. *Midterm student assessment*

7. Performance Assessment of Statistical Problem Solving

Although not typically used by teachers in their individual classes, the detailed scoring of open-ended statistical problems is now being used in the Advanced Placement Statistics Examination in the United States. A scoring rubric is used by test scorers to evaluate students' statistical knowledge and communication skills, using a scale of 0 to 4 for each scale (Olsen, 1998). A score of 4 on statistical knowledge means that the student completely understands a problem's statistical components, synthesizes a correct relationship among these components, uses appropriate and correctly executed statistical techniques, and provides reasonably correct answers. A score of 4 on the communication scale means that the student's explanation of what was done was clear, complete, organized, and correct, that appropriate assumptions and caveats are stated, diagrams or plots used are appropriate, and an appropriate and complete conclusion is stated.

The different assessment methods described above may be used in combination with each other as well in combination with traditional quizzes and exams. Chance (1997) provides an excellent model for combining different assessment components.

5 An Assessment Framework

In considering different possible assessment approaches, questions emerge as to when to use a particular method, which method works best in a particular situation, and how to best utilize assessment results. An assessment framework can be used to help make decisions such as these (see Garfield, 1994). The framework is helpful in planning assessments in the courses as well as in thinking about evaluating curricular projects.

The components of the framework are:

1. *What to assess*: concepts, e.g., skills, applications, attitudes, and beliefs.
2. *The purpose of assessment*: e.g., why the information is being gathered and how the information will be used.
3. *Who will do the assessment*: e.g., the student, peers, or the teacher.
4. *The method to be used*: e.g., quiz, report, group project, individual project, essay, or another method.
5. *The action to be taken and the feedback given to students*: e.g., suggestions for revision and improvement, a summative grade, topics to review, etc.

The multiple components of this framework demonstrate that assessment is a complex system with a major role in a statistics course, not just "writing a good test". This framework challenges educators to consider why they are using an assessment, how it will be used, and how it will be evaluated; in order to select or develop the most appropriate method.

6 Practical Implications for College Statistics Teachers

It is important that instructors of statistics courses carefully think through and state in their course syllabi learning objectives for students, and then check to see that the assessment methods used adequate reflect these objectives. Students quickly learn that what is assessed is what is valued. Therefore, it is important to make sure that assessments focus on the knowledge and skills viewed as most important for students to learn. For example, if a course goal is to enable students to be able to read, interpret, and critique statistical concepts and graphs used in the media, a student assessment should be described in the course syllabus that is designed to determine if students have developed these competencies. As assessment methods are revised in light of course goals, some words of caution are important to consider.

One caution relates to the introduction of student projects as a method of assessing student learning. The assessment literature includes recommendations to provide exemplars, or models of what the instructor considers to be a good or excellent job on a performance task. This is viewed as helpful in communicating the instructor's standards and in enabling students to develop similar standards. One challenge in providing exemplars (which students eagerly examine) is that can it lead to conformity and inhibit creativity. Students typically want their projects to look just like the good examples in order to earn a high grade. Therefore, it is important for instructors to develop ways to provide a variety of exemplars with commentary that will provide guidance and high standards but also prevent students from providing carbon copies of the exemplars provided.

Another caution relates to trying to do too many new things at one time. Statistics educators eager to experiment with new forms of assessment may be overwhelmed and discouraged if they abandon the methods they are familiar with and instead embrace an entirely new set of techniques. Instructors trying alternative methods of assessment may want to keep in mind that each assessment activity is a way to provide students with feedback on how to improve their learning and not just as an activity used to assign a grade. Rather than trying to incorporate many new assessment methods in their course at one time (e.g., introducing weekly minute papers, students projects, and portfolios), they might feel more successful (and less overwhelmed) if instead they pick one new method, try it out for a while, and then gradually introduce and experiment with other techniques. This is especially important because instructors need to allow time to reflect on the assessment information gathered from students, and to monitor the impact of these results on their perceptions of the class and their teaching.

The use of alternative assessment methods, particularly those involving papers and reports, often results in overwhelming amounts of paperwork to read and evaluate. Even with carefully constructed scoring rubrics, it takes time to evaluate and provide detailed feedback to students. Each time a new task is assigned, the instructor needs to be cognizant that the paper must be carefully read and good and timely feedback provided to students. It is much easier to have a teaching assistant score a multiple-choice test. However, students learn more from the challenging tasks of practicing their communication skills and integrating and applying their statistical knowledge, which they do when completing a project or written report.

A final caution deals with the courage sometimes required to administer an assessment to students and then carefully examine the responses. The results may reveal that instructors' perceptions of how well they have taught something, or how well students have learned something, may be incorrect. For example, when using minute papers in class, students might be asked to explain (in writing)

their understanding of a particular concept (e.g., the standard error of the mean, a standard deviation, or what a sampling distribution represents). Asking the students to describe their understanding in these anonymous minute papers is easy. Reading their statements takes courage and patience. The often awkwardly-worded responses sometimes reveal a serious lack of understanding, despite the brilliance of an instructor's explanation or the most carefully designed in-class activity.

7 Current Assessment Challenges

Now that many statistics educators are developing and experimenting with alternative assessment methods, some unique challenges have emerged. Gal & Garfield (1997) describe some of these challenges:

1. Assessment of Students in Computer-assisted Environments

This involves two separate subtopics: searching for effective ways to assess what students can do and how they reason when they use computers or other technological aids, and the nature of and limitations on inferences that can be drawn from assessments when students learn with computers but are tested without computers, as is common in many classes.

2. Assessment of "Statistical Literacy"

This involves testing the application or transfer of student learning to interpretive or functional tasks such as those encountered in media or outside the classroom. The challenge in assessment of statistical literacy is that it should involve examining not only what students can do and how they think when asked to, for example, reflect on a report in the media, but also their tendency or disposition to do so without being cued.

3. Assessment of Students' Understanding of "Big Ideas"

Throughout their education students encounter important statistical ideas such as variation, error, bias, sampling, or representativeness. These ideas underlie much of students' overall understanding of the uses and limitations of statistics, yet may not have direct mathematical or functional representation. Their meaning may depend on the context in which they are invoked. Tasks are needed that can assess students' understanding of and sensitivity to the prevalence and importance of such "big ideas" in different contexts.

4. Assessment of Students' Intuitions and Reasoning Involving Probability Concepts and Processes

This area has seen much research activity and there are clear indications that many students have misconceptions or intuitive beliefs that are not being changed during instruction (e.g., Konold, 1995). Research documents that students may compute probabilities correctly on quizzes yet appear to misunderstand the big ideas. For example, students may indicate that they know that the means of small samples are more likely to deviate from the population mean than means of large samples. Yet, when given a problem like the one shown in figure 3, many reveal that they do not really understand this idea at all. The data accompanying this problem were gathered from two large samples of university students at the end of an introductory statistics course. One sample is from the USA and one is from Taiwan. This is part of a comparative study that reveals the persistence of misconceptions related to probability, chance, and random events for students in both countries.

Figure 4 shows another item that illustrates students' misunderstanding of probability, referred to

Half of all newborns are girls and half are boys. Hospital A records an average of 50 births a day. Hospital B records an average of 10 births a day. On a particular day, which hospital is more likely to record 80% or more female births?

	Taiwan	USA
a. Hospital A (with 50 births a day)	8.6%	9.7%
b. Hospital B (with 10 births a day)	30.7%	37.9%*
c. The two hospitals are equally likely to record such an event	60.7%	52.4%

* correct answer

Figure 3. *Item assessing misconception about small sample variability*

When two dice are simultaneously thrown it is possible that one of the following two results occurs:

Result 1: A 5 and 6 are obtained (regardless of order).
Result 2: A 5 is obtained twice.

Select the response that you agree with the most:

	Taiwan	USA
a. The chances of obtaining each of these results is equal	73.0%	63.7%
b. There is more chance of obtaining result 1	20.5%	27.7%*
c. There is more chance of obtaining result 2	2.5%	4.5%
d. It is impossible to give an answer. (Please explain why)	4.0%	4.1%

* correct answer

Figure 4. *Item showing the equiprobability bias*

as the equiprobability bias (Lecoutre, 1992). Less than one-third of the students in each sample gave the correct response to this question, revealing what researchers have labeled "an equiprobability bias", or the tendency to view all outcomes as equally likely. Research suggests that if students had been asked to compute each probability separately, they might have computed the correct responses, and if prodded, have then selected the correct response.

There is a need to transfer and adapt promising assessment methods and instruments used by researchers (mostly involving in-depth clinical interviews with selected students) to formats that are reasonably acceptable and accessible to teachers and that can be used for "routine" classroom use, such as the two items shown in figures 3 and 4.

5. Assessment of Outcomes of Group Work

The assessment and grading of students' work when it is done in groups has been described as a frequent stumbling block for statistics teachers who are considering incorporating cooperative group activities and projects in their classes. There are concerns about grading fairness and the challenge

of motivating all students to participate equally.

6. Developing Models to Use in Evaluating and Comparing Curricula

As new curricula, innovative textbooks, and instructional software replace traditional approaches to teaching statistics, there is an increasing need for reliable, valid, practical, and accessible assessment instruments to use in evaluating the relative utility of these materials and methods. As long as statistics items used in large-scale or standardized assessments remain focused on computations (as opposed to statistical reasoning) and provide little context, the relative effectiveness of statistics courses or units will remain difficult to ascertain.

7. Using Assessment to Determine What Students Understand after They Interact with Simulation Software Packages

Although many statistics instructors share a belief that students need to see sampling distributions created on the computer, using dynamic representations that vary population characteristics and sample size, few instructors explore just what students actually learn from these activities. delMas, Garfield & Chance (1997) have been using pre and post test items to assess students' understanding of sampling after interaction with a simulation program. They found some surprising results, such as: students were attending to unimportant details (e.g., the height of the bars) and continuing to have misconceptions about the shapes of sampling distributions from non-normal populations (e.g., the shape of sampling distributions based on small samples should represent the shape of the population). By examining students' responses to items like these, they have been challenged to change the way they use the simulation software with students.

8 Summary

The assessment revolution has contributed heavily to the reform of statistics education, offering new ways to better understand and document student learning. It is important for all statistics instructors to become familiar with the new assessment concepts, procedures, and methods, and to carefully determine how they may best evaluate student learning and assess the outcomes for their particular courses, keeping in mind that the most important role of assessment is to enhance student learning.

References

Angelo, T. & Cross, K.P. (1993). *Classroom Assessment Techniques*. San Francisco: Jossey-Bass.

Birenbaum, M. & Dochy, F. (Eds.) (1996). *Alternatives in Assessment of Achievements, Learning Processes and Prior Knowledge*. Boston: Kluwer Academic Publishers.

Chance, B. (1997). Experiences with authentic assessment techniques in an introductory statistics course. *Journal of Statistics Education* [Online] **5** (3).

Cobb, G. W. (1998). *The objective-format question in statistics: Dead horse, old bath water, or overlooked baby?* Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Darling-Hammond, L., Ancess, J. & Falk, B. (1995). *Authentic Assessment in Action: Studies of Schools and Students at Work*. New York: Teachers College Press.

delMas, R., Garfield, J. & Chance, B. (1997). *Assessing the effects of a computer microworld on statistical reasoning*. Paper presented at the Joint Statistical Meetings, Anaheim, CA. August.

Gal, I. & Garfield, J. (Eds.) (1997) *The Assessment Challenge in Statistics Education*. Amsterdam: IOS Press.

Garfield, J. (1994). Beyond testing and grading: using assessment to improve students' learning. *Journal of Statistics Education* [Online] **2** (1).

Hubbard, R. (1997). Assessment and the process of learning statistics. *Journal of Statistics Education*, [Online] **5** (1).

Jolliffe, F. (1997). Issues in constructing assessment instruments for the classroom. In *The Assessment Challenge in Statistics Education*, Eds. Gal & Garfield. Amsterdam: IOS Press.

Keeler, C. (1997). Portfolio assessment in graduate level statistics courses. In *The Assessment Challenge in Statistics Education*, Eds. Gal & Garfield. Amsterdam: IOS Press.

Konold, C. (1995). Issues in assessing conceptual understanding in probability and statistics. *Journal of Statistics Education* [Online] **3** (1).

Lecoutre, M. P. (1992). Cognitive models and problem spaces in "purely random" situations. *Educational Studies in Mathematics* **23**, 557–568.

Mackisack, M. (1994). What is the use of experiments conducted by statistics students? *Journal of Statistics Education* [Online] **2** (1).

Moore, D. (1997). New pedagogy and new content: the case of statistics. *International Statistical Review* **65** (2), 123–137.

Olsen, C. (1998). *What's a rubric? Scoring the open ended questions*. Paper presented at the Joint Statistical Meetings, Dallas, TX.

Schau, C. & Mattern, N. (1997). Assessing students' connected understanding of statistical relationships. In *The Assessment Challenge in Statistics Education*, Eds. Gal & Garfield. Amsterdam: IOS Press.

Starkings, S. (1997). Assessing student projects. In *The Assessment Challenge in Statistics Education*, Eds. Gal & Garfield. Amsterdam: IOS Press.

Watson, J. (1997). Assessing statistical thinking using the media. In *The Assessment Challenge in Statistics Education*, Eds. Gal & Garfield. Amsterdam: IOS Press.

Wild, C., Triggs, C. & Pfannkuch, M. (1997). Assessment on a budget: using traditional methods imaginatively. In *The Assessment Challenge in Statistics Education*, Eds. Gal & Garfield. Amsterdam: IOS Press.

Résumé

La contribution entreprend une déscription de l'interaction entre des objectifs nouveaux envisagés pour les étudiants et des méthodes alternatives pour évaluer leur apprentissement réel. Elle offre des suggestions aux enseignants des statistiques désirant de reexaminer leurs propres pratiques d'évaluation en classe aux lumières de cette évolution, en donnant pour exemples quelques approches innovatifs utilisés dans des courses introductoires en statistiques et en decrivant des demandes actuelles adressées aux enseignants.

[*Received February 1998, accepted October 1998*]