ALIGNING ASSESSMENT WITH INSTRUCTIONAL GOALS AND VISIONS

Iddo Gal, University of Haifa, Israel
Joan B. Garfield, University of Minnesota, USA

This paper outlines eight common goals for introductory statistics education and examines assessment issues and dilemmas related to these goals. It argues that increased attention should be placed in both instruction and assessment on the use of processes and tasks that enable students to generate, express, and explain opinions and develop communication skills. The paper outlines some challenges for statistics educators who seek to align assessment practices with specific instructional goals and broader visions.

INSTRUCTIONAL AND ASSESSMENT VISIONS

Efforts to rethink traditional approaches to college statistics education have been represented in recent years in a growing stream of journal articles and conference proceedings (Moore, Cobb, Garfield, and Meeker, 1995). Changes in content and pedagogy are being recommended as part of a “statistics reform” effort directed at both the college and pre-college levels. While attention to the teaching of statistics has become visible, with few exceptions (e.g., Gal and Garfield, 1997; Hubbard, 1997), almost no publication has addressed *assessment* issues in statistics education. This is despite the pivotal role that assessment plays in informing instruction and students’ motivation, and despite extensive research showing that statistical concepts are often poorly understood.

The essence of good educational assessment is that it is fully aligned with teaching goals. A recent report by the Mathematical Sciences Educational Board (1993) has enumerated three principles of good assessment in the mathematical sciences. These principles are repeated below with the word “mathematics” replaced with “statistics”:

The Content Principle: Assessment should reflect the statistics that is most important for students to learn.
The Learning Principle: Assessment should enhance statistics learning and support good instructional practices.
The Equity Principle: Assessment should support every student’s opportunity to learn important statistics.

Focusing on the first principle, we ask, What are the “Important Statistics,” the target concepts, knowledge, understanding, and dispositions that are to be learned, and therefore, to be assessed? The answer is not straightforward, as introductory statistics curricula may differ in their orientation (e.g., theoretical vs. applied), level of mathematical sophistication, extent of coverage of probability and inferential topics, or degree of reliance on computers. Such diversity noted, our working assumption is that the majority of students encounter statistics in only one or two courses, either a generic introductory service course, or in courses couched within the needs and applications of a specific field, such as psychology or engineering. We further assume that most students are more likely to be *consumers* of data rather than full-fledged researchers. With this in mind, a two-part overarching vision emerges for the outcomes of basic statistics education (see Gal, Ginsburg and Schau, 1997). After finishing their introductory encounter with statistics,

A. Students should be able to make sense of statistical information prevalent in an information-laden society, and be sensitive to and comprehend instances and processes of variability and uncertainty in the world around them,
B. Students should be able to contribute to or take part in the production, interpretation, and communication of data and findings pertaining to data-based problems they encounter in their professional or civic life.

We have identified eight interrelated basic goals for statistics instruction that can support achieving this broad vision, from relevant literature on mathematics, statistics, and science education and from our own prior work. Following the presentation of each goal, we briefly comment on some points or dilemmas that have implications for assessment.

SPECIFIC INSTRUCTIONAL GOALS AND ASSESSMENT ASPECTS

Goal 1: Understand the purpose and logic of statistical investigations. As a background to any deep understanding of statistics, students should understand the big ideas that underlie data-based inquiries, and develop a sense for *why*, in what *circumstances*, and under what *assumptions* statistical investigations are initiated and conducted. Examples for such ideas are:

a. The need to describe populations or target groups by collecting data
b. The presence of variation; assumptions about “normality” of distributions
c. The need to study and to infer from samples to populations
d. The need to “reduce” raw data by using summary statistics and displays
e. The notions of error in measurement and inference, and the need to control and estimate such errors

Comments: It is unclear to what extent “big ideas” are addressed and in turn assessed in introductory statistics courses. Some of them may seem too conceptual to instructors focused on teaching students how to "do" statistics, or may be covered in separate courses on research methods. Either way, teachers may prefer to assess specific skills..

Goal 2: Understand the process and phases of statistical investigations. This goal is the one most commonly listed by teachers and textbooks as the organizing principle for instruction in introductory statistics courses. It suggests that students should understand the actions and considerations involved in each phase of a statistical inquiry::

a. Formulate a question;
b. Plan the study (e.g., approach and overall design, sampling, choice of tools);
c. Collect data and organize it;
d. Display, explore, and analyze data
e. Interpret findings in light of the research questions;
f. Discus interpretations, conclusions, conjectures, and implications.

Comments: Some teachers simulate these stages by working on fictitious or given datasets, to focus on "doing" statistics and save time on data collection. Other teachers engage students in data-analysis projects focused on research questions the students care about, collect their own data, and prepare summary reports. This makes the interpretive phase more meaningful, and allows for a genuine audience to whom a report has to be presented. This goal is central to the acquisition of statistical knowledge and skills, yet *how* it is being realized in the classroom has implications for generative and interpretive skills that can be developed and for assessments that can and should be employed.

Goals 3 and 4: Master procedural skills and Understand mathematical relationships. Students have to master the component skills needed in the analytic stages of a statistical investigation, such as being able to organize data in relevant formats, construct useful data displays, or compute needed indices by hand or assisted by technology. Also, students have to understand, intuitively and (sometimes) formally, key underlying mathematical ideas, e.g., how the mean is influenced by extreme values in a dataset, or how variability within and between groups is reflected in the values in a T-test.

Comments: It is important to find if students can perform certain procedures, and perhaps explain the mathematical underpinnings of some procedures. Yet, we should just as much want to evaluate if students recognize *when* and *under what circumstances* certain statistical procedures have to be used, *why* they have to be used, *what* is their advantage over other procedures, and *what interpretations* can be attached to their results.

Such knowledge is hardly captured by forced-choice techniques, and is best assessed by examining what students actually do on open-ended problems.

Goal 5: Understand probability and chance. Some informal grasp of probabilistic ideas underlies the understanding of the logic behind topics such as sampling or tests of significance and is needed to follow the reasoning of statistical inference. A common belief is that experiences with chance behaviors starting with random generating devices (e.g., coins) and leading to (computer) simulations can help students to develop intuitive understanding of key ideas such as that a model can simulate events and help estimate real-world probabilities, or that statistical tools (such as tables in the back of statistical textbooks) serve to estimate probabilities under certain assumptions.

Comments: Equiprobability events cover only a subset of the uncertainties and chance processes involved in real-world events or systems. As has been repeatedly shown (Shaughnessy, 1992), students' intuitions and knowledge regarding chance processes may be patchy even after formal instruction about basic laws of probability. What may seem like "understanding" of formal rules (e.g., knowing the likelihood of certain outcomes of dice rolling) may provide little indication as to "understanding" in another domain involving probability (e.g., inference from samples to populations).

Goal 6: Develop interpretive skills and statistical literacy. Sense-making skills should be developed as part of work towards the previous goals, mainly the ability to interpret basic data patterns (e.g., frequency listings, tables), to understand the meaning of results from basic statistical procedures, and to present opinions about the overall meaning of multiple pieces of evidence, in light of specific research questions. "Statistical literacy" is used here to highlight the expectation that most students who study introductory statistics will seldom have to collect or analyze their own data after graduation, but instead will need to be able to make sense of material reported in the media (e.g., results from surveys or opinion polls). They will have to comprehend and evaluate the acceptability of claims presented in concise formats that may combine text with summary statistics or displays, and that provide partial information about samples, methods, or tools used..

Comments: In both "interpretive" and "statistical literacy" contexts, students should develop an appreciation for possible limitations on the conclusions that can be drawn from data due to, e.g., reliability of measurements, representativeness of samples, experimental design. The assessment of students' opinions about data or text-based media items is not a trivial matter (see below). Further, teachers may assume that students who can provide acceptable interpretations of their own data context can transfer their interpretive skills to statistical literacy contexts, yet this claim has not been subjected to systematic research.

Goal 7: Develop the ability to communicate statistically. Students should be able to effectively communicate, orally or in writing, about a range of experiences, technical issues, and results that come up during investigative processes. This may include discussing the degree of confidence that can be attached to reported findings, arguing thoughtfully about the validity of other people's interpretations of results, or phrasing questions, such as about acceptability of generalizations from a certain study.

Comments: This is a "process" goal in the service of the preceding six goals. It is important. It is not meant to be assessed as a separate outcome of learning statistics, but is subsumed by assessing progress towards the other goals.

Goal 8: Develop useful statistical dispositions. We want students to,

a. *Hold positive beliefs and attitudes*: Certain beliefs, attitudes, and self-perceptions are a precondition for effective learning in a statistics class yet also important outcomes of participation in it (Gal, Ginsburg, and Schau, 1997). Students should depart with a view of themselves as individuals capable of statistical and probabilistic reasoning, and with a sense of comfort, willingness, and interest to think statistically.
b. *Adopt a critical stance*: A questioning attitude is useful when faced with arguments or reports that purport to be based on data.
c. *Appreciate the power of statistics*: Recognize that planned statistical research can provide information that may assist in making personal, social, and business related decisions in the face of uncertainty, and may lead to better conclusions than relying on anecdotal data or on intuitions (but that this is not guaranteed).

Comments: Given that quite a few students enter statistics courses with "anxiety," it seems useful to monitor certain feelings, attitudes, and beliefs students may hold before and during a class, to detect those students who may need assistance. Yet, the entities to be assessed are "slippery" and it is difficult to assess them reliably.

IMPLICATIONS FOR INSTRUCTION AND ASSESSMENT

The eight goals above illuminate areas that need further attention in introductory statistics courses, in terms of both instruction and assessment. These goals deal with diverse aspects of the process of doing statistics (goals 2-5), yet emphasize the importance of deep understanding of foundational ideas (goal 1), and aim to increase attention to the need to develop sense-making, interpretive, and communicative capabilities, and a reflective approach to statistics (goals 6-8).

We need to adopt assessment methods that can provide richer information about the degree of progress towards each of the eight goals outlined above, but also about the degree of *integration* between students' skills, knowledge and dispositions. Such integration can be revealed, for example, when students have to manage meaningful data-analysis situations that require them to function both as planners, generators, and interpreters of data and findings, or when asked to critically evaluate a media report.

One underlying theme in the above discussion is that students should increasingly develop an ability to render reasoned *opinions,* which may be in the form of descriptions, judgments, explanations, or assumptions, and are especially central when results have to be interpreted and generalizations, inferences, conjectures, or conclusions reached in connection with the questions that motivated a study. Yet, opinions about findings very often cannot be characterized as "right or wrong", but instead have to be evaluated in terms of their degree of reasonableness, in light of the quality of the evidence available and procedures (e.g., sampling, tools) used to generate this evidence.

We argue that adequate assessment of learning outcomes related to all goals outlined above is not possible using only traditional formats such as multiple-choice or brief structured questions, despite their seemingly high reliability and ease of use. Such methods are too often divorced from context. They focus on correct application of formulas or accuracy of graphs or computations, thus assessing only one or two of all learning goals (but see Wild, Triggs, and Pfannkuch, 1997), and cannot assess opinions very well.

A credible determination of the reasonableness of students opinions about data, however, cannot in many cases be made without having access to both the opinions *and* to the arguments on which they are based. A greater emphasis on communicative processes and methods is thus required in both instruction and assessment. Communicative tasks may require students, for examples, to report about data-analysis projects (Holmes, 1997) in written forms (full report, poster, short newsletter article); prepare oral presentations;

keep process summaries (project logs, team journals; see, e.g., Stromberg and Ramanathan, 1996); or analyze real-world media items and write critiques or reflective pieces (such as a letter to a public official following the reading of a statistically-oriented article).

The need to adopt new assessment methods that can encompass progress towards all learning goals forces statistics educators to struggle with the relative emphasis they place on reliability versus validity of the different kinds of information they can collect. The choice of assessment methods should be aligned with specific instructional goals and geared to serve our overall instructional vision, to increase students' ability to contribute to and effectively function in an information-laden, statistically-oriented society.

REFERENCES

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

Gal, I., Ginsburg, L., and Schau, C. (1997). Monitoring attitudes and beliefs in statistics education. In I. Gal and J. B. Garfield (Eds.), *The assessment challenge in statistics education* (pp. 37-51). Amsterdam: IOS Press.

Holmes, P. Assessing project work by external examiners. (1997). In I. Gal and J. B. Garfield (Eds.), *The Assessment Challenge in Statistics Education* (pp. 153-164). Amsterdam: IOS Press.

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

Mathematical Sciences Education Board (1993). *Measuring what counts: A policy brief*. Washington, DC: National Academy Press.

Moore, D. S., Cobb, G. W., Garfield, J., and Meeker, W. Q. (1995). Statistics education Fin de Siecle. *The American Statistician, 49*(3), 250-260.

Moore, D. S. (1990). Uncertainty. In L. A. Steen (Ed.), *On the shoulders of giants: New approaches to numeracy* (pp. 95-137). Washington, DC: National Academy Press.

Shaughnessy, J. M. (1992). Research on probability and statistics: Reflections and directions. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 465-494). NY: Macmillan.

Stromberg, A. J. and Ramanathan, S. (1996). Easy implementation of writing in introductory Statistics Courses. *The American Statistician, 50*(2), 159-163.

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