

From Gal, I. & Garfield, J. B. (editors). The Assessment Challenge in Statistics Education. IOS Press, 1997 (on behalf of the ISI). Pages 17-25. ISBN 90 5199 333 1. Copyright holder: International Statistical Institute. Permission granted by ISI to the IASE to make the chapters from this book freely available on the Internet. The References section is a separate chapter and is available on the Internet at <http://www.stat.auckland.ac.nz/~iase/publications/assessbkref>.

2. Some Emerging Influences Underpinning Assessment in Statistics

Andy Begg

Purpose

A number of influences impinge on statistics education. This chapter focuses on three of these that are especially noticeable at the K-12 level but also operate to some extent at the college level: a) the changing place of statistics in the curriculum, b) the emphasis on processes in the mathematics curriculum, and c) new ideas about learning and the ways that assessment is viewed in education in general. These, together with the purposes and principles of assessment, are explored because they underpin aspects of assessment in statistics education.

STATISTICS IN THE CURRICULUM

The first influence is the changing emphasis on statistics. It is used in many disciplines and is needed for participation in society, hence it is gaining prominence in education. This is evident from the emergence of statistics departments in universities and curriculum initiatives in schools that incorporate statistics and suggest that its teaching should begin at an early stage (NCTM, 1989; Australian Education Council, 1991). These curricula emphasise *statistics and probability for all*, which contrasts with the past focus on mathematically mature students in high school and college or those who needed statistics for work in other disciplines.

Traditionally curricula have been subject-based. Now curriculum frameworks are being developed (Ministry of Education, 1993) where two other dimensions are considered: the *essential learning areas* and the *essential skills*. Learning areas (language, science, technology, mathematics, social sciences, arts, health) may be thought of as subjects but are intended to provide a framework for all that is taught in schools. If one considers a matrix with learning areas juxtaposed against school subjects, one can see how the subject mathematics contributes to learning areas (science, technology, social sciences, ...) and how subjects (chemistry, geography, economics, industrial arts, home economics, ...) contribute to the learning area mathematics. With this perspective, when we talk of statistics within mathematics we need to know whether we are talking about the subject or the learning area. The former gives a single-subject focus, while the latter implies cross-curriculum consideration which may be acceptable to teachers of young children but may present a challenge to high school teachers.

A. Begg

Similarly, the essential skills (related to communication, numeracy, information-handling, problem-solving, self-management and cooperation, work and study) can be considered against learning areas or subjects. These skills are generic and can be learned through all subjects, including statistics. They suggest an approach to assessment that emphasises interpretation and communication, the use of computers to access and process data, problem-solving, understanding concepts rather than computation of results, and working with others rather than working individually.

MATHEMATICAL PROCESSES

The second emerging influence is evident from recent curriculum documents (NCTM, 1989; Australian Education Council, 1991; Ministry of Education, 1992). It is the increasing emphasis on mathematical processes (what mathematicians do) alongside content (what mathematicians know), and these processes link closely with the essential skills from curriculum frameworks. In statistics these processes relate to what statisticians do. One listing of the processes is problem solving, reasoning, communicating, making connections, and using tools. These processes are intended to be integrated with content and not treated as separate topics.

Assessment should similarly be concerned with this meld of content and processes. The assessment implied by these processes involves using extended and open-ended tasks that allow students to investigate meaningful problems, reason about data, communicate verbally and write reports, make connections and models, and use computers. It is likely to involve formal and informal observation of oral and written work so that evidence can be collected regarding the quality of the statistical thinking, the understanding of the use of tools, and the connections between situations and their models. To indicate the influence of these processes on assessment, some aspects of each are listed.

Problem solving

Problem solving is more than solving word problems; it includes investigating, modelling and simulating, and doing open-ended project work. It may lead to extended practical projects and assessment needs to reflect this. Some aspects of this are considered by Lesh, Amit, and Schorr in Chapter 6, by Holmes in Chapter 12, and by Starkings in Chapter 11.

In assessing problem solving one would look not only for solutions but for evidence of planning (identifying and exploring problems; formulating plans; collecting, retrieving and analysing data), using a range of strategies (making tables, using graphs, finding relationships), and modelling (using concrete materials, graphs, networks, tree diagrams and flow diagrams; then fitting a model to the data and checking its fit).

Reasoning

Mathematical reasoning is concerned with proof and certainty within given assumptions, but when mathematics moved to problems with multiple solutions this degree of certainty weakened. Now that statistical and probabilistic thinking has been introduced, there are further aspects of uncertainty in reasoning that require consideration.

2. Some Emerging Influences

Aspects one might look for in assessing reasoning include classifying and describing (sorting and organizing data to support an argument); arguing (justifying answers and procedures, drawing conclusions from interpretations, using summary statistics appropriately); inferring (making and evaluating conjectures and hypotheses, interpolating and extrapolating, making appropriate and responsible decisions); and proving (acknowledging assumptions, constructing proofs).

When reasoning is valued the focus for assessment moves from the final answer to the way it was obtained. Ideas about models and the statistical reasoning linked to them are developed by Lesh, Amit, and Schorr in Chapter 6; and by Batanero, Godino and Navarro-Pelayo in Chapter 18.

Communicating

Statisticians are consulted by clients because of their statistical expertise, but their knowledge within a client's field is often limited. Statistical consulting involves finding what the client wants and becoming aware of what extra data may be useful. After collecting and processing all the relevant data, the task is to interpret the results and convey the findings to the client in non-technical terms. Thus, communication is fundamental in statistics and should be reflected in its teaching and assessment.

Communication is also important because it influences the success of teaching and learning programs. It occurs at the inter-personal level (working cooperatively, debating possible courses of action) and at the intra-personal level (understanding what needs to be done, reflecting on and clarifying thinking). It involves listening and speaking (discussing difficulties, asking questions, presenting and explaining results, discussing the implications and accuracy of conclusions, discussing possible interpretations of data, presenting reports); reading and writing (using reference material, writing a report); and representing (using graphs, diagrams, and symbols to represent ideas). While some of these aspects are generic and will be assessed in other subjects, they may also need to be considered within statistics.

Some aspects of communication are considered in more detail, including those related to statistical literacy (Watson, Chapter 9) and projects (Holmes, Chapter 12; and Starkings, Chapter 11).

Making connections

Making connections is important for successful learning and because statistics pervades life and is used in many subjects. Aspects that need to be connected occur within the subject (linking concepts, procedures, and topics within statistics and probability; relating various representations of concepts and procedures; recognising equivalent representations; seeing the subject as integrated), with other curriculum areas (linking statistics with other subjects and with generic skills), and in everyday life (using statistics in work, leisure activities and in familiar and unfamiliar situations; understanding statistical representations in the media). These connections need to be made explicitly, and this linking needs emphasis through encouragement and by being recognised in assessment.

Making connections with real-life problems is considered by Lesh, Amit, and Schorr in Chapter 6; connecting procedures and concepts is discussed in Chapter 7 by Sloane, Kelly and

A. Begg

Whittaker; and making links between statistical concepts is a focus in Chapter 8 by Schau and Mattern.

Using tools

Tools are used to gather and process data, to help make sense of them, and to simulate situations. Tools allow computation to be de-emphasised, but if they are used in class then they need to be available during assessment. Elements that one might look for in the use of tools include general uses such as measurement in the collection of data, calculator use for computation, and computer use with applications (word processing, databases, spreadsheets, packages for statistics and graphing, and simulations).

The use of computers in assessment in statistics is discussed further by Lajoie in Chapter 14 and by Cohen and Chechile in Chapter 19.

CONSTRUCTIVISM

The third emerging influence on assessment is from new theories of learning. This issue has been addressed in numerous studies and conferences on assessment in mathematics. For example, in an ICMI study Romberg (1993) gave a general overview of learning, and Galbraith (1993) reported on some of the problems fitting assessment with constructivism. Leder (1992) discussed some similar issues related to constructivism.

Until recently assessment was based on behaviourism. The assumption was that knowledge was broken into small and specific objectives that were transmitted to students, and each objective could be assessed. Behaviourist assessment provided feedback so teachers could make decisions about programs and report on their students. This assessment was usually after learning (summative), and prior knowledge was not seen as important. Some teachers did informal diagnostic (before learning) and formative (during learning) assessment, but because their programs were well defined this did not usually influence their teaching. The focus for behaviourist assessment was on correct responses to questions and on behavioural objectives which usually related to the repetition of taught procedures or the recall of facts. Little consideration was given to understanding concepts, to applying ideas to other subjects and new situations, or to modelling, investigating, or communicating.

The recent move from behaviourism to constructivism requires a consideration of different assumptions. Von Glasersfeld, whose ideas have been accepted by many mathematics educators, assumes that the basic principle about constructivist learning is *knowledge is not passively received but actively built up by learners*; and that *coming to know is an adaptive process in which learners organise their experiences of the world, rather than discover an objective reality* (1989). Teachers influenced by constructivism need to consider what feedback they and their students want from assessment. This feedback may answer a number of questions which arise before, during, and after teaching (Begg, 1991).

Before teaching the questions include:

2. Some Emerging Influences

- What are the student's interests, ideas, conceptions and misconceptions about the related content and processes prior to teaching?
- What are their questions about the topic likely to be?
- What activities could focus on their questions?

These questions suggest informal assessment strategies such as classroom discussion, stimulating students to ask questions, and observing and listening to students at work. Some of the answers to these questions will be known by experienced teachers and the information may be partially gained from teacher-sharing sessions.

During learning the questions include:

- What are the students wanting to know about the topic?
- What processes are they tending to use?
- Are the learning activities focusing on these processes?
- Are the meanings being constructed similar to the intended ones?
- How are the students putting together their ideas?
- Are they developing learning-to-learn skills?
- How might the unit of work be modified to better suit the class?

These questions require the teacher to find out what students are thinking while they are learning. This is likely to be done informally during teacher-student interactions when the teacher is supporting, encouraging, and challenging with “what if” and “what if not” questions that respect the learners' autonomy and their views, but help them move from their incorrect ideas.

Developing learning-to-learn skills suggests that students need strategies to identify what they are learning from the experiences they encounter and to build their ability and confidence with self-assessment. Empowering students to audit their own learning involves helping them to recognise criteria that point to development in learning, teaching strategies for reflection, inviting them to compare their self-assessment with the teacher's assessment, and assisting them in comparing their present knowledge (knowing that, knowing about, and knowing how) with that of others so they can appreciate that other views are possible.

After a topic is covered the assessment questions are:

- What are the students' ideas now and what processes have they been using?
- Are their present ideas and skills different from their earlier ones?
- Can their new ideas and process skills be used in unfamiliar situations?
- What needs to be reported or documented?
- What changes need to be made to the program?

“After learning,” or summative assessment, is difficult because it is multi-faceted. It becomes even more difficult with a negotiated curriculum, because, even though activities are chosen by the teacher with specific objectives in mind, these activities are only starting points for open-ended approaches and negotiation with the students. These negotiations determine the

A. Begg

direction of the lesson and therefore what learning occurs. Testing for standard concepts will not show all that has been learned in this situation; even in traditional classes it is likely to lead to teaching for tests rather than teaching for understanding. In spite of these difficulties, the teacher needs to know what has been learned and may find this by observation, by marking work, by formal assessment events, or by means such as discussion and concept maps.

PURPOSES FOR AND TYPES OF ASSESSMENT

While the implications of constructivism suggest numerous purposes for assessment, these can be grouped in an overview as two basic purposes: a) for better learning by providing feedback to teacher and learner, and b) for reporting to parents and in-school documentation. It is reasonable to assume that if assessment does not contribute to at least one of these, then it would be better not to assess. Even when it does contribute, if it is time consuming, then one must ensure that its usefulness outweighs the value gained by using the time on learning.

Assessment for better learning occurs mainly before and during teaching. It involves finding what students bring to their learning, whether they are ready for new learning, and what their interests and ideas are. Sometimes assessment for better learning is to provide motivation by building students' self-esteem and by changing their perception of their teacher's expectations of them; in this case if marks are allocated, then they need to be clustered around a high mean.

Assessment for reporting and documentation occurs after teaching. School documentation is important to facilitate the movement of students who change schools and classes and who leave school, and it is needed so that schools can report to parents. These forms of assessment and methods of reporting are influenced by community expectations.

While these two purposes of assessment are generally accepted, the types of assessment that should be used, especially after learning, are more problematic. Behaviourist assessment has usually been norm-referenced with a student's performance compared with that of peers, or criterion-referenced with performance measured against well-defined specific objectives. Wiliam (1994) is typical of educators who see limits with these types, and he suggests that we need to consider two further types—ipsative and construct-referenced assessment. Ipsative assessment is concerned with "value added," that is, the difference between what was known before learning and what is known after. Construct-referenced assessment relates to holistic and open-ended activities such as projects, where it may not be appropriate to define specific objectives or standards. With this type of assessment different students may investigate different topics, but a general set of standards exist from which the assessor selects the appropriate ones for each student's work.

While ipsative and construct-referenced assessment suggest a movement from behavioural assessment, these four types are interrelated. Underlying any criteria are normative assumptions made about what learners of a particular age might generally be expected to know. Underlying "value-added" assessment is the need to find out what was known before and after, and this is likely to involve criteria. Criterion-referenced and construct-referenced assessment relate to external standards, the difference being whether these relate to explicit objectives or whether they involve a choice of objectives for each particular student's work.

PRINCIPLES OF ASSESSMENT

2. Some Emerging Influences

Regardless of the types of assessment used, an agreement by educators should guide the use of the principles of assessment. These principles relate to consistency, focus, range of techniques, individuality, self-assessment, reporting, openness, and practicality.

Consistency

Assessment should be consistent with its purposes (which need to be explicit), and with one's view of statistics and learning. Tasks should yield valid results about learning, and what is assessed should reflect what is important in statistics. What is important is not only facts and skills but includes the processes previously discussed. Thus assessment should involve a range of tasks and require the application of a number of ideas rather than relying on tests which focus on narrow sets of skills such as the correct application of standard algorithms. Assessment should promote equity and avoid inequities such as those related to gender, ethnicity, language and culture.

Focus

Assessment should focus on what students know and can do rather than on what they do not know or cannot do (Cockcroft, 1982), and on the process of coming to know. This suggests tasks should be accessible at numerous levels because all students will not be working at the same level. The tasks also should be more than computational so that alternative conceptions and students' thinking are identified.

Accepting a range of abilities implies having a range of achievement levels. These could be hierarchical—definitions, basic facts, and standard procedures at the lower level; making connections and problem solving next; and statistical thinking and reasoning, communication, and generalisation at the highest level. A preferred approach uses three categories of goals—content, process, and thinking goals, because all students will achieve in some way in each area.

Range of techniques

The limitations of written tasks with students who are not good readers or writers but have verbal facility are obvious, but less obvious are problems with multiple-choice questions where some learners may be more inclined to guess correct answers while other learners may reflect more carefully on options and, with good reasons, think of other factors that make the correct solution problematic. To overcome such limitations a range of techniques is needed. This may include informal tasks (questioning, observing, listening, reading students' writing), written tasks (multiple-choice, matching lists, short answer, completing statements, open-ended questions, short essays, open-book tests, two-stage tasks), oral tasks (presentations, seminars, interviews, debates, role plays, interviews), practical tasks (investigations, projects, simulations, experiments, use of computers and calculators), and cooperative tasks (group activities, peer assessment). A selection of such tasks could form the basis for an assessment portfolio.

A. Begg




Individuality

Assessment should allow for the unusual because the unique aspects in students' responses often help show what sense they have made of the experiences provided and what strategies they have used. It is not unusual for a student to discern a pattern, devise a strategy, or come up with an idea that has merit, which the teacher has not encountered. For a teacher there is fascination and reward in detecting the unusual. An example of this is where a student assumes that a probability is conditional while the teacher assumes it is not, and the student's thinking needs to be probed before the response can be assessed.

Self-assessment

Self-management skills and the aim of establishing an ethos of life-long learning have implications for self-assessment. Teachers can facilitate this by providing lists of topics to be covered, understandings to be gained, an indication of time for various tasks, the criteria being used for assessment, and opportunities to self-assess with marking guides.

Students may use learning logs to record learning, effort, and enjoyment. Such logs may be confidential or be read by the teacher but this needs to be agreed upon at the start. Young children and older students can describe what they have learned from an experience and what they would like to explore. This is particularly useful if there is an emphasis on learners establishing their prior ideas about a topic, identifying their strengths and weaknesses, considering the ideas and strategies of others, and monitoring their own changes. Learning logs may be used to stimulate reflection. Within logs self-assessment sheets might be used and be headed, "Things I have found out," "Things I now know," and "Actions I took to sort out my problems." A section for students' remaining questions may be included and be headed, "Things that still puzzle me" or "Things I would still like to know." For younger children "talking heads" might be used:

		
I think I understand it	I am coming to grips with it	I need help!

Reporting

Reporting provides students and parents with an indication of progress, and Cockcroft (1982) makes it clear that "a cross is of little assistance to a pupil unless accompanied by an indication of where the mistake has occurred and an explanation of what is wrong or a request to consult the teacher when the work is returned." Similarly a pass/fail, a grade, a percentage mark, or a ranked position does not give the same amount of information as a sentence or two written about the student.

The student's right to privacy needs to be considered; some countries have a "privacy" act to ensure this. Feedback for students and parents should not be public. This means that teacher feedback should be by written comment or by a quiet word rather than said openly in class.

2. Some Emerging Influences

Openness

The assessment process should be open to review and scrutiny by students and parents. Students should know what is expected, what criteria are used, when deadlines occur, and how results will be used. If a student misses an assessment deadline then the teacher should give a “best estimate” of the student’s work, and not a zero as this is unlikely to stand scrutiny in terms of what the student knows.

Practicality

The assessment program must be practical. It should not be too time consuming for the students or for the teachers, yet it needs to provide information for the teacher to improve the learning program and feedback for students and parents.

IMPLICATIONS

The focus of this chapter was not to discuss assessment alternatives as such, but to make explicit a basis on which assessment can be built. From the issues considered a pattern emerges which supports the approaches introduced in later chapters. These issues fit with ideas from the National Council of Teachers of Mathematics (1995) in the Assessment Standards and are advocated by writers such as Stenmark (1989, 1991), who introduced many teachers to authentic assessment, performance assessment including investigations and writing, observations and interviews, and the use of portfolios. The approaches fit with statistics education because the holistic nature of the subject is taken into account, content and processes are assessed, feedback needed to implement programs based on constructivism is produced, and the approaches fit with the purposes and principles of assessment. These alternatives are discussed in the chapters that follow, in particular, authentic assessment is discussed in Chapter 3 by Vos and Colvin, written tasks by Joliffe in Chapter 15, and portfolio assessment by Keeler in Chapter 13.

Bringing about change is not easy, and if teachers are to change, they require time, support, and encouragement. The first step is for teachers to see assessment in statistics education as problematic and to identify it as an area of their concern. They may start by thinking about any of the issues discussed; alternatively, they may start with a dissatisfaction with their present practice and be looking for specific techniques, but hopefully at some stage these teachers will distance themselves from the specifics and consider the ideas that underpin them.