

1

Assessment and feedback in statistics

Neville Davies and John Marriott

1.1 Introduction

Statistics is not only a subject in its own right but is also applied to diverse other subjects, including the sciences, geography, psychology, business and economics. Consequently assessment may need to cover a very wide-ranging set of topics and activities taught within a statistics curriculum. In this chapter we consider the implications of the ubiquity of the application of statistics and its pedagogy, the significant impact of this broad 'learning base' on what should be assessed and the plethora of ways assessment and feedback can be provided.

1.2 Types and purposes of assessment

There are four types of assessment that we identify in relation to statistics:

1. *Diagnostic assessment* seeks to identify the starting position of students, to identify gaps and thus enables these to be filled.
2. *Formative Assessment* seeks to use assessment for improvement, to indicate strengths and weaknesses and to give both student and teacher insight into the progress being made; formative assessment can also contribute some marks to overall assessment.
3. *Summative assessment* seeks to evaluate overall achievement, usually at the end of the course. With mid-course summative assessment it is possible to use the resulting feedback for formative purposes.

4. *Evaluative assessment* is the final putting together of results and it involves the whole assessment of the programme of study. Its aims include the satisfaction of society, the institution, the teachers and, hopefully, the candidate. Such assessment covers the quality assessment of the course itself.

It is common sense that good assessment will lead to both good learning and good teaching, thereby encouraging a balanced view of the subject. MacGillivray argues in Chapter 2, the case that ‘assessment is for learning’ and that this is promoted by variety in statistics assessment. In the past there has been an emphasis on summative assessment. There is, however, an increasing recognition that only having assessment at the end of a course in statistics helps neither students nor staff. An increasingly varied approach to assessment has evolved in recent years, together with emphasis on feedback to students and staff. Bingham (2001) shows how assessment fits in with the whole teaching process. He clearly shows the interrelationships between learning outcomes, strategies for learning and teaching, assessment strategies and criteria for assessment, arguing cogently that assessment is an integral part of the whole process of teaching and learning.

Statistics, a complex subject with many components, provides the paradigm and methodology for making sense of data arising from a wide range of subjects. This means that we need to think long and hard about how we assess and provide feedback to students of statistics. The key reasons for doing assessment include enabling both students and their teachers to realise what they know, the teachers to know their own effectiveness, the institution to award grades to students, and society to judge the effectiveness of the awarding institution. No one, simple technique is going to provide all of this information. Thus, we take a broad view of the many methods and approaches available for assessment and feedback, but emphasise those most relevant for the statistics teacher.

1.3 What are we assessing?

Ideally an examiner will have been involved in designing the learning outcomes and assessment criteria of the course documentation. The assessment would then be constructed to determine reliably whether the learning outcomes have been met. In a statistical context there will inevitably be a range of topics requiring assessment. These follow from the fact that statistics as a subject, provides a paradigm for problem solving and is a methodology for doing science and many other subjects. Thus, the assessment of statistics should include assessment of:

- Understanding the statistical problem-solving approach (PSA).
- The course content, for example technical aspects of statistics.
- The process of doing statistics, for example undertaking an analysis that may include estimating parameters.
- The use of global skills, for example the use of appropriate statistical software.

- Personal and transferable (key) skills, for example skills inherent in statistical problem solving.
- Critical abilities such as those required in looking at the practical aspects of surveys.
- Communication of statistical results and conclusions.

In addition, the learning outcomes and their assessment need to take note of the levels at which learning in general and statistical learning in particular operate, and ensure a reasonable coverage.

Before looking at the positive aspects of assessment it is worth noting some of the problems that can occur in carrying it out. Brown (2001) provides a useful list of these problems, which includes: mismatches between the stated outcomes and the work required of the students; unclear or unstated assessment criteria; overuse of a particular form of assessment; and issues concerning feedback to students.

1.4 Formative assessment

Formative assessment of the student's work in statistics enables appropriate action to be taken by both student and teacher to correct misunderstandings. Black and Wiliam (1998) showed that effective feedback was a major means of developing learning. For this reason formative assessment is essential for students' involvement, motivation and understanding of their own progress and ways forward.

Successful formative assessment will employ 'resource-light' methods yet provide reliable information. Wider use of online testing and feedback has opened new ways of dealing with such assessment and some universities now make significant use of computer-based testing and feedback. Examples of these include online diagnostic testing and broader based tests that progress through a course with different levels of difficulty. Formative evaluation is often informal, but it can be formalised, for example by using short student-marked tests in class time or past papers for revision.

There is the problem that sometimes only the better students undertake work that does not count for their examination 'final' marks. Some statistics courses cope with this by awarding a very small proportion of the summative marks for all work done, including the formative assessment. Another approach is to get students to keep a log of all their formative work, some of which is later randomly selected for summative assessment. Sometimes marks are given for just doing the quiz or assessment, irrespective of the results. Logbooks are particularly useful for assessing the practice of statistics.

Marks, in general, may not be appropriate in formative assessment. Although good marks may give some encouragement, anecdotal evidence suggests that marks can detract students from looking carefully at their work. Marks give a sense of closure, whereas the purpose of formative assessment is to enable

students to fill in the gaps and deepen their understanding. Hence discursive feedback comprising, for example, comments on suggested reading and reference to other class work, is essential.

There needs to be a relationship of trust between student and teacher and the formative evaluation must give a constructive and honest judgement of the work done. To be most effective, formative evaluation must be ongoing and continuous, with a bias towards the early parts of a course. The performance of the class as a whole can also inform the teacher of class weaknesses and strengths and so provide some overall feedback.

The teacher needs to encourage self-evaluation and provide resources or strategies to aid this. Providing self-tests that students can do repeatedly to check their improvement in understanding can be of use here.

General questions presented to the whole class to provide feedback to the teachers are often done in an ad hoc fashion, on the spur of the moment. However, it is better to plan this activity carefully. One consideration is to ask questions that deliberately probe different levels of understanding.

Whatever technique is used there is a need to encourage responsible self-assessment by students. There needs to be a parallel commitment to the value of formative assessment by the teacher, a sharing of learning outcomes and a confidence that all students have the ability to develop and improve in statistics.

1.5 Feedback

The two directions of feedback are from the lecturer to the students and from the students to the lecturer. Both are necessary for effective learning and teaching. The students need to know how they are progressing and lecturers need to know the effectiveness of their teaching and of the students' learning. For each of these there is general feedback about overall issues of class progress and progress through the curriculum. There is also specific feedback on the individual student's understanding of individual topics and the teacher's approach to them. Feedback, which can apply to the teacher as much as to the student, has been shown to be a key element in the learning process. Snelgar and Maguire in Chapter 3 discuss the use of feedback on very early assessments in a staged approach to quantitative projects.

Formative assessments of stages in a course, a central element of the feedback, are often supplemented by the use of general questionnaires to obtain feedback from students. Many universities carry out regular surveys of student opinion and experience, the outcomes of which can give useful broad views but they need to be supplemented at both programme and individual course level. To get further detail, an end-of-course questionnaire is often used. In one course students were asked to rate the following topics on a five-point scale: quality of teaching; formative feedback; handouts; organisation; and opportunity to achieve and contribute. Feedback from such end-of-course questionnaires is useful for planning the next run but it is too late for the students concerned, so it should be

supplemented by mid-course feedback. In general it is important that at each level of detail some feedback is obtained before too much learning has taken place. The longer that misunderstandings and misconceptions in statistics are left, the harder they are to deal with.

Possible tools for rapid feedback from the students that are readily available are:

- Short tests to explore understanding from a previous class.
- Minute papers with questions such as, ‘What part of today’s work was most difficult to understand?’
- Wall sheets, placed near the exit, for students to write comments on as they leave the class.

At the very end of a course it is worth spending somewhat longer with groups of students noting and making suggestions with regard to its good and bad points, the teaching and their operation as a class. To ensure that the feedback is effective and that the teacher has the time to respond, it is useful for the teacher to keep a logbook of comments. This will enable an overall picture to be developed and suitable changes initiated. The logbook should also be used at times of marking to record topics that students have found difficult or easy and to note any problems and misconceptions that become evident.

This topic has so far been discussed in the context of the feedback that is of value in improving the quality of teaching. However, the teacher should also review the appropriateness, structure and order of the topics in the curriculum. Should topic X still be taught in the same way at the same time when the course is next delivered? In a review of the curriculum, one needs feedback from a fresh range of sources. These sources should include:

- Current research.
- General literature on the development of statistics.
- General literature on what topics are finding practical application.
- The work experience of past students.
- Employers.

Statistics has some specific issues that need to be addressed. For example, for statistics within a service, or non-specialist course, the students’ perceptions of the link between statistics and their main area of study need careful exploration. Such feedback can help in the choice of illustrative material later in the course.

Feedback to students tends to be at an individual level. It is often helpful to give comments on the overall results of the class – common strengths, weaknesses and misconceptions. This approach is also of value when giving end-of-course feedback to students who will often go on to use their learning, and benefit from encouragement and support at this stage.

Feedback to students on their progress should include:

- The precise knowledge, understanding and skill that they should possess at the current stage of a course.
- The level of knowledge, understanding and skill that they have actually reached, and what the gaps are.
- How to go about closing these gaps.

Feedback frequently tends to focus on the second of these elements. One approach that goes some way towards addressing all three is to start each class with a short, student-marked test on the work covered in the previous class. Anecdotal evidence suggests that the most effective way of improving standards of learning is to improve the quality of feedback provided to students – we propose that more research be conducted into this area.

1.6 Summative assessment

Summative assessment needs to be carried out against the formally stated learning outcomes and assessment criteria of a course. There should be a clear match ensuring that all learning outcomes are covered. This may need an imaginative use of assessment methods. If it is felt that some learning outcome is not and cannot be covered, then the learning outcomes need to be re-specified.

The characteristics of summative evaluation are that:

- It is formal, in the sense of being official and to be taken seriously.
- It gives marks or grades.
- It may have several components, but between them they represent the whole of the course.
- It often contributes to the grade and so requires particular care from the assessor.
- It often occurs towards the end of the course.

There are many ways of carrying out summative assessment, with alternatives to consider within each method.

1.6.1 The examination paper

Traditional examination papers are based on a selection of questions to be done in a fixed time. Issues of importance are the comparability of level of questions in a paper, course coverage that is equitable and fair, and the need to ensure that the weaker students have a chance to pass and that the brighter students are challenged. This can be achieved by using structured questions, with sections that are progressively more demanding. It is then essential to ensure that getting

the first part wrong is not a total barrier to progressing further with the question. Assessors and moderators need to look at this aspect of papers and examination boards need to examine the mark distributions in papers to assess issues of fairness. The classic approach has been to ask a question on standard theory at the beginning of a question and then ask for it to be developed or applied in the latter parts. This can lead to an excessive emphasis on ‘bookwork’ and to the mistake of assuming that such bookwork can be classified as ‘easy’. Many applications, developments and scenarios can be either segmented into, or introduced by, a progression from easy to more difficult elements.

In statistical data analysis and inference courses, and also in courses on probability and statistical modelling, questions and even whole papers use the analysis of data as the means of testing understanding of statistical ideas and processes. One type of question that is typical of this approach gives the background to a real problem and a copy of computer output of various analyses, and then asks for comments or a report, depending on the balance within the paper. Such questions can sometimes provide a direct link to the student’s specialisation in a service course. As mentioned before, care is needed to avoid complicating the statistics by contexts that are not yet sufficiently familiar to the students. Examiners should avoid making the statistics more difficult by assuming knowledge that the student might not have; cultural and language issues must also be given careful consideration. Another means of making links, or of providing a direct application for a specialist student, is the design of questions round a pre-read research paper. Other forms of question include: essays on general topics; questions requiring the design of surveys or experiments; the construction, fitting or criticism of models; and, for the more mathematical course, the proof of theoretical results.

1.6.2 Using multiple-choice questions

Multiple-choice questions have advantages in that they are relatively quick for students to answer, so the range of ideas tested can be wider than in a traditional examination. They can also be marked more easily, by using a grid or by using computer marking. They take thought and care to set, as the ‘wrong’, or distracting, answers play an educational role in the testing. Depending on the question, they need to be as plausible as the ‘right’ answers and/or embody common mistakes. Ideally all distracters should be equally ‘wrong’. There is also the limitation of needing questions that have clear and specific answers. However, this ‘limitation’ is the downside of the necessity for the examiners to think very carefully about what they want the students to know and do, which is a positive benefit of this approach to assessment. A style that can be useful in statistics is to ask students whether particular statements are appropriate or inappropriate. This is useful in assessing students’ understanding of such aspects as interpreting plots and graphs or computer output from analyses. An important use of multiple-choice testing is for formative assessment. Here interest is not in the right–wrong decision but in the nature of the misunderstanding that led to a particular wrong answer. An aspect of multiple-choice questions that needs

consideration, and is less restrictive than often thought, is the marking scheme. Multiple-choice questions do not have to be equally weighted, provided students know what the ‘marks’ are.

One great advantage of feedback obtained in this way is that if online computer methods are used the assessments can be taken and marked very quickly. This makes quick regular checks on students’ understanding possible for the timely provision of support. A further advantage available in many computer-based testing systems is that each student can be given data randomised for a given model and then marked accordingly. This clearly reduces problems of plagiarism.

It may be noted that studies of computer-based assessment find, on the whole, no significant difference in students’ assessed levels between paper-based assessments and their computer-based equivalents. See Thelwall (1999, 2000) and the references therein. In both approaches, what matters is the quality of the question design and of the marking.

1.6.3 Short answer, or comment questions

As well as the commonly used ‘tick the correct answer’ responses, a multiple-choice question environment can be used to elicit short phrases or comments from students. Such phrases are most efficiently received and assessed via a virtual learning environment. Multiple-choice questions usually involve the student scrutinising statements made by the teacher, whereas it is sometimes more appropriate, and may give better assessment coverage of a course, if a teacher tried to assess how students might respond to a statistical statement in their own words. This approach can be particularly useful in data analysis where a scenario and computer output are provided but, instead of asking for a report under examination conditions, a series of questions requiring only a sentence or phrase response are asked. This goes beyond the limitations of providing answers for multiple-choice questions, without requiring report writing under examination conditions and time restrictions.

1.6.4 The individual assignment

There are several advantages in the use of assignments. They can face students with open-ended problems and with issues of problem definition, and also enable a variety of ideas and methods to be combined and applied in a real context. The role of individual assignments is discussed in detail by other authors later in the book.

1.6.5 The group assignment

An important aspect of the statistician’s role is that of being a team member. Therefore the assessment needs to encourage and measure this role on tasks

suitable for group work. These include tasks already mentioned as suitable for individual assignments, but involving more complex issues and data sets. To these we can add consultancy for a client, e.g. design of a quality control system for a small local manufacturer, or the mini-project type of activity where there are several sub-tasks to be completed. Classic examples are the survey and the experiment, where there are issues of design, piloting, implementation, data collection, data analysis, reaching practical conclusions, presentation and reporting. Examples of group assignments using car price data and agricultural data are given by Mashhoudy in Chapter 21 and by Lopez *et al.* in Chapter 14.

As with all forms of assessment, the first task of the teacher is to develop clear criteria for the marking of assignments. The next problem is the allocation of marks against criteria. In the case of the group assignment there is, firstly, the breakdown of marks for the sub-tasks, with the flexibility to allow for the fact that different teams may adopt differing approaches, or indeed may be carrying out different studies. Secondly, there is the question of group versus individual marks. Common options are:

- All members of a team get the same mark, which reflects the practicalities of team-work, but may be unfair if one member does not contribute.
- A team mark contributing, say, 80% and an individual mark contributing the remainder. This latter mark may be awarded by the assessor on the basis of some individual task within the assignment or on the basis of interview or observation; alternatively, the mark may be allocated by the team itself, for example by allowing each member to spread a fixed total of marks around the team and then averaging.
- A team mark modified by a multiplicative factor representing the contribution of the individual. See, for example, Goldfinch (1994).

One approach to team reporting and assessment is the poster display project discussed by Griffiths and Sheppard in Chapter 4. Here, the team members, perhaps with the use of one or two large posters, seek to show the problem and their solution in a concise form. This demands clear thinking and is an excellent means of sharing ideas and experience between teams and with a wider group of staff. A set of criteria can be devised for the assessment of the poster display, which produces a team mark to go with individual marks.

1.6.6 Continuous assessment

There are many methods of continuous assessment. Regular written tests provide one method: to relieve pressure and give room for formative assessment, the students can submit a selection of these as a portfolio for the summative assessment. An alternative is to use assignments instead of a final assessment. These may be marked in stages as a form of continuous assessment.

1.6.7 The portfolio

Any course on statistics that emphasises the processes of statistics will seek to assess a wide range of different aspects of the students' learning. This assessment will need to include the student's own constructed knowledge and evidence of their learning skills. The development of a portfolio by the student brings together not just the tasks undertaken during the course but also the student's reflection on them, possibly via some form of journal. It is helpful for the students to be given sample questions to ask of themselves, such as: How did you respond to the feedback obtained from an assessment? How would you improve your work on the task done? What relationships do you see between Topic X and Topic Y? How could the operation of the group on the previous mini-project have been improved?

1.6.8 Self-assessment

Learning theory emphasises the importance of students, both individually and as part of a group, taking responsibility for their own learning. An important part of this is self-assessment. This covers a broad spectrum, from a general sense of progress on a topic to the self-marking of tutorial questions. It also includes students working in pairs or groups to mark each others' work. This type of activity helps in the overall process of self-assessment and is increasingly used by staff, both for educational motives and to help them focus on their supportive role as teachers.

1.7 Techniques and practicalities of assessment

When a teacher is designing an assessment careful consideration should be given to:

- The learning outcomes and the capabilities/skills (implicit or explicit) they imply.
- Methods of assessment that match with these outcomes and skills.
- The relative efficiency of different methods in terms of student time and staff time.
- The advantages and disadvantages of any proposed method.
- The forms of marking scheme or criteria that are appropriate.

This approach should naturally lead to the design of a specific assessment task, but care must be taken to avoid common pitfalls. These include: unintentional ambiguities in a question or assignment, miscalculation of the time and resources required to do the assignment or to mark it, and failure to design either a suitable set of criteria or a marking scheme.

The design of effective assessment tasks in statistics can be time-consuming. Essay questions are quick to set but lengthy and subjective to mark; to be reliable, very detailed marking schemes are required. Mathematical statistics questions often take a long time to set but are fairly quick and objective to mark. Statistics questions in general can be lengthy to set and lengthy to mark, and need careful marking schemes.

It can be instructive to consider questions set by colleagues in other subject disciplines. There will almost certainly be questions that are intriguing and surprising, and also forms of question that may be useful for assessing statistics. Occasionally the content, as well as the form of question, can be relevant to statistics (Brown, 2001).

In considering the assessment of different types of activity, great care needs to be given to defining the process for allocating marks. It is sometimes helpful to have two sets of criteria, one dealing with the specifics of the activity, the other setting out general descriptive criteria for deciding what ranges of marks are appropriate. The marking seeks to reach consistent answers for the two sets of criteria.

Before discussing any methods in detail it is important to underline the role of assessment. Many students view assessment as merely a test of memory after having done all the work. However, in the areas discussed here assessment is an integral part of the learning experience. It ensures that the students carry out the processes of statistics and is thus essentially formative as well as being summative. This fact indicates that more attention needs to be given to feedback to students, even where the assessment provides summative results.

A vital element of statistical assessment is the assessment of students doing statistics. Some of the choices facing the teacher are illustrated below.

1.7.1 Tutorial exercises and examples

The classic student experience of doing statistics is that of doing tutorial problems. A common practice is to ensure that students can start on the problems before the tutorial so that some tutorial time can be dedicated as necessary to clearing up difficulties students identify and discussing the harder problems. Making best use of tutorial time in this way means that well-structured tutorial sheets need to be available to students at least one week before the tutorial session. An informal record-keeping system will enable this approach to be used more effectively for formative assessment.

1.7.2 Data studies

Most students will use their statistics to investigate and analyse data. Thus, most courses will involve studies of sets of data. To be effective the study needs to be well focused. Most data can lead into a range of different lines of study. Students need to be prevented from spending time on peripheral issues and so they should

be assisted in understanding some of the issues of choice between alternative lines of enquiry.

For service courses data drawn from the main discipline is clearly a good source for these students and will also interest them. The complexity of the study will depend on the tools available for any analysis, in particular the software available.

1.7.3 Practical work in the classroom

Rather than being supplied with data, it can be a valuable experience for students to collect the data for analysis themselves, in specified classroom experiments. The potential pitfalls in conducting these types of experiments have led to such practical work rarely being used for summative assessment. This is not a major problem in formative assessment, for which such practical work is commonly used, though time pressures for staff and students seem to have reduced the use of practical work actually carried out in a classroom environment.

1.7.4 Investigations

The availability of data on the Internet provides a vast resource for students to carry out investigations into problems of relevance to their main study or of current interest. Ideally an investigation will be centred on a real problem. In such investigations student assessment will probably be based on three elements:

- Ability to use statistical knowledge.
- Strategies adopted to solve the problem.
- Ability to communicate and report on the problem and its solution to the problem owners.

1.7.5 Consultancy

For postgraduate study, many US universities have an ‘intern programme’. This provides a university-wide consulting service using staff and students. Some UK universities have used student teams to provide consultancy to local businesses and industries, or within the university. Assessment may be based on both written and verbal interactions with, and presentations to, the client. Consultancies can also be simulated, with a staff member – preferably from another department – acting as the client and the work centring on a scenario with data/information revealed gradually, as the student asks for it.

1.7.6 Case studies

Case studies provide an application-oriented form for teaching. The assessment tends to take the form of reports on the case to the ‘problem owners’. This promotes key problem-solving skills, but can miss the assessment of detailed

technical knowledge. One way of tackling this is to generate artificial and specific problems that test knowledge, but which are set within the context of the case study. As usual it is essential that students have detailed information on the marking methodology of all assessments within the case study.

1.7.7 Mini-projects

Mini-projects can be built around finding information on a specific topic or methodology. They can also be part of a major project, such as fitting a defined model to a set of data from the major project. The assessment here would follow that used in a major project. Due to the mini-nature of the work, the assessment would tend to be of the final output rather than of the process.

1.7.8 Projects with dissertations

For final-year undergraduate and MSc projects, the assessment tends to be based on the final output, but the time allocated also allows some consideration of the assessment of the process of the research. This may require a detailed research strategy and project plans after the first month of work or take the form of mid-way presentations on work in progress or course poster sessions on the research in the latter stages.

1.7.9 Statistical problem-solving

Unfortunately, as Marriott *et al.* (2009) show, assessing statistical problem-solving needs far more thought than at first meets the eye. They demonstrate that teaching and learning statistical problem-solving requires a paradigm shift in teaching and that the cognitive skills expected of students are not what are traditionally expected. It follows that assessing statistical problem-solving is a much more complex procedure than is achieved by assessing through traditional examinations. More research is needed in this area.

With so many approaches, the teacher is clearly faced with frequent problems of choice. One guide in this choice can be found in Anderson and Krathwohl (2001) who provide a taxonomy for teaching, learning and assessment. The authors consider the interaction between categories of what they refer to as the cognitive and knowledge dimensions. An important aspect of their approach is the careful examination of the verbs used in stated teaching and learning outcomes, and how they relate to categories of the cognitive process. Their Table 5.1 provides the detail and a helpful list of verbs.

Anderson and Krathwohl use a table with rows and columns that are the categories of the knowledge dimension and the cognitive process dimension respectively, to show how a teacher might approach the assessment of learning outcomes. In order to illustrate this, consider the learning outcome (LO): 'The students should learn to use confidence intervals for inference'.

In Table 1.1 we provide a taxonomy for this LO.

Table 1.1 A taxonomy table for the assessment of confidence intervals.

		Cognitive Process Dimension					
		1 Remember	2 Understand	3 Apply	4 Analyse	5 Evaluate	6 Create
Knowledge Dimension	A Factual knowledge						
	B Conceptual knowledge		TL	LO TL	TL	TL	
	C Procedural knowledge	EX		LO TL EX		TL	
	D Metacognitive knowledge	TL		TL			

For the above LO, the verb ‘use’ means ‘implementing’ in this context, which is associated with ‘Apply’ in Anderson and Krathwohl’s Table 5.1. In order to accomplish the ability to ‘use’ successfully, the students must identify the type of problem they are faced with, decide on and select the inferential procedure that is required (the appropriate confidence interval), and then apply the appropriate procedure to compute the interval. If we now look at the verbs that have been used to describe the tasks the students must undertake (identify, decide, select and apply) they would indicate that the implementing in this example requires both ‘Conceptual knowledge’ and ‘Procedural knowledge’. Cells B3 and C3 in Table 1.1 are therefore associated with the learning outcome in this case and are marked with ‘LO’. In teaching to this outcome the student would be encouraged to use ‘Metacognitive knowledge’ by being taught how to assess how they know the problem fits into a particular type. Then they would ask themselves whether the answers they arrive at are ‘sensible’ at the end of the process of constructing a confidence interval. The students are also taught to ‘Evaluate’ their answer in context. The cells B2, B3, B4, B5, C3, C5, D1 and D3, marked ‘TL’ in Table 1.1,

are all associated with the teaching and learning needed to support this specific learning outcome.

In planning an assessment, the teacher must now choose appropriate assessment(s). A formative assessment may well be appropriate if the knowledge and cognitive processes identified in the consideration of the teaching and learning supporting the specific outcome are to be addressed. If a strict assessment of the learning outcome is required, the assessment could be tailored to assessing the types of knowledge and cognitive processes associated with cells B3 and C3 alone, for example: ‘Calculate the 95% confidence limits for the large sample confidence interval given your sample data’; to answer this question, the student must ‘Remember’ and ‘Apply Procedural knowledge’. (The associated cells in Table 1.1 are marked ‘EX’.) It can be seen that this question only aligns with one of the cells identified as applying to the stated learning outcome.

1.8 Assessment of active learning approaches

Individual assessments are supported by detailed marking schemes leading to percentage marks or grades. These are often supplemented by descriptive assessment criteria that link the student’s overall performance on the item of assessment to the more general learning outcomes. The two should provide consistent appraisals. Any inconsistency needs careful appraisal by teachers and Examination Boards. Often, two staff will mark a project and agree an overall mark.

Table 1.2 presents an illustrative set of words suitable for assessment criteria for different types of learning outcomes. Three categories of Fail, Pass and Distinction are used. Some criteria seek to subdivide pass into 40–49, 50–59 and 60–69. The detailed marking scheme should normally do this.

One example of an approach to designing a marking scheme and then allocating marks for assessment is provided by Marriott *et al.* (2009). They adapt the delightfully simple grading scheme for UK Advanced Level coursework in statistics provided by the curriculum development organisation Mathematics in Education and Industry (MEI). The scheme allocates the assessment questions to domains for grading and uses a very simple mark allocation scheme suggested by Garfield (1994, Example 2). Allocated marks correspond to the responses candidates make to each question being incorrect (0 marks), partially correct (1 mark) or correct (2 marks).

1.9 Conclusions: assessment strategy – principles and guidelines

In planning for success in assessment, it is useful to take on board some core principles and follow succinct guidelines.

- *Principle 1:* Content – assessment should reflect the course content that is the most important for students to learn.

Table 1.2 Assessment criteria for different learning outcomes.

Form of Learning Outcome	Fail	Pass	Distinction
Knowledge of the theory of least squares.	Shows no knowledge of theory of least squares.	Appropriate range of knowledge of theory of least squares.	Theory of least squares used critically and innovatively.
Comprehension: Understand the method of least squares as applied to linear regression.	Significant lack of understanding, misunderstanding of both method and regression.	Adequate to competent comprehension of least squares and of its application.	Explanation comprehensive, accurate and insightful.
Application of the method of least squares.	Poor and erroneous use of method.	Largely correct and appropriate use of method.	Accurate and insightful use of method.
Analysis of a real problem using a data set.	Few ideas and lack of recognition of context.	A reasoned set of analyses set in the problem context.	A comprehensive and insightful set of analyses, clear conclusions for the context.
Synthesis (i) Carry out a library-based project on a new area of application and produce a report in survey form.	Little apparent understanding of the area and an unstructured and confused report.	Good review and coverage with clear, balanced and structured explanation. Well referenced.	An insightful and critical report, Excellent coverage of the literature with an integrated and concise structuring.
Synthesis (ii) Design and carry out a small survey.	Design inappropriate to objectives. Poorly implemented.	An appropriate design competently carried through.	Design and implementation show depth of insight and thoroughness of implementation.
Evaluate a published study involving the statistical analysis of medical data.	Inadequate appreciation of the problem, the data and the analysis.	Correct understanding of the problem and methods, with a correct identification of the strengths and weaknesses of the methods used.	Shows critical insights into the issues and methods involved with the correct and imaginative advocacy of possible alternatives.

- *Principle 2:* Process – assessment should enhance the sound development of concepts, insights and deeper ways of thinking; it should encourage deep learning.
- *Guideline 1:* Students value what is tested, so test what you value.
- *Guideline 2:* Doing formative assessment implies that the student wants to know as well as wanting to pass.
- *Guideline 3:* Good summative assessments will seek to ensure that the student understands as well as passes and that all passing students have reached a basic competency.
- *Guideline 4:* If formative and summative assessments do not match then generally the summative will dominate, as students are generally more interested in passing than in understanding.

These six principles and guidelines require that the stated LOs and assessment criteria clearly reflect what is valued by the students, are relevant, and are important for students to understand.

A final consideration in deciding on an assessment strategy is the sheer practicality of carrying out the marking involved. Consistency of marking is important, and one way of ensuring this is to have more than one person marking each item of assessment. It is also important to remove any possibility of personal prejudice in marking; one way of doing this is to use anonymous marking. It is possible to develop efficient techniques that fit the statistical content, the type and range of students and the requirements of the learning outcomes. Some of these often raise concerns as being too time-consuming. It should be noted that, with experience and ingenuity, most forms of assessment can be carried out effectively and efficiently. In essence the teacher's best assessment strategy is to choose the best balance of methods to achieve the given ends.

