® LEARNING AND ASSESSMENT IN STATISTICS

PETOCZ, Peter & REID, Anna
Macquarie University
Australia

The main role of assessment is to support learning, and any view of assessment implies a corresponding view of learning. Research on students' conceptions of statistics, learning in statistics and assessment, suggests that there is a clear variation from narrow to broad views. Another dimension is students' perceptions of their future professional roles and how that impacts on their present studies. In order to support the learning process, assessment should be structured in such a way as to make apparent to students the full range of variation in conceptions and to encourage them towards the broadest and most inclusive ideas. Further, it is important that the approach to assessment has coherence with the overall pedagogical approach.

INTRODUCTION

In *Learning to Teach in Higher Education*, Ramsden (1992, p.5) claimed that: *"The aim of teaching is simple: it is to make student learning possible."* First and foremost, the role of assessment is to support the process of student learning and to help prepare them for their future professional roles. This view subsumes the traditional focus on checking that students have learned things, which gives teachers an awareness of what their students haven't learned and hence allows them (hopefully) to modify their teaching appropriately. It also includes the role of assessment for certification or for evaluation of the teaching program itself. And further, any scheme of assessment, or even individual assessment task, implies a particular view or theory of learning, and hence should be coherent with it.

A decade ago, an interesting and influential collection of chapters was published (Gal & Garfield, 1997) on the challenge of carrying out assessment in statistics. The editors put forward a pragmatic approach to the challenge, first listing what is to be assessed, then outlining the (then) current challenges: doing versus using statistics, the role of mathematics in statistics and the need for alternative approaches to assessment. They also pointed towards the 'future challenges' of assessment: in computer-assisted environments, of statistical literacy, of understanding of 'big ideas', of intuition and reasoning about probability, and of outcomes of group work. They summarised the challenge as: "*to identify assessment methods that are able to elicit and reveal student learning corresponding to each of the eight subgoals …*" (p.7).

The book contains chapters on theoretical, practical and novel aspects of assessment in statistics representing various complementary (and sometimes contradictory) viewpoints. The book is now available on the IASE website as an historical record of the state of thinking a decade ago, and valuable source of ideas for current use.

Some of the 'future challenges' identified by Gal and Garfield were already addressed in chapters of the book (e.g., assessment of statistical thinking, Watson, 1997; assessing students in groups, Curcio & Artzt, 1997) and they continue to be the subject of writing since that time. As well as including examples of assessment of specific topics in specific situations (for example, Lunsford *et al.*, 2006; Pange, 2006), the recent statistics education literature contains much about assessment of statistical literacy, probabilistic reasoning and the 'big ideas' of statistics. For instance, Garfield and Ben-Zvi (2005) outline an epistemological model of variability and give suggestions for assessing students' "deep understanding" of each of the components of variability. Chance (2002) examines the nature of statistical thinking and its assessment and gives creative examples of 'big picture' examination questions, while delMas *et al*. (2006) report on the development of a bank of assessment items to assess statistical reasoning (the ARTIST project). MacGillivray (2006) discusses assessment – including group assessment – of students' ideas of probabilistic reasoning. Sowey (2006) addresses the problem of showing students that statistics is worth studying by formulating challenging questions that can be used for assessment (or self-assessment) during a statistics course, allowing students to observe "its resilience to challenging questioning": for instance, *"What would statistics be like if there wasn't the Central Limit Theorem?"* Budé (2006) summarises the debate concerning assessment of statistical literacy, reasoning and thinking, recommending assessment of statistical understanding using three levels of questions: definitions and procedures at a superficial level, inference from data at an intermediate level and applications using transfer of knowledge at the expert level.

In general, much of the writing on assessment in statistics stays within the specific confines of statistics education rather than making explicit use of the broader literature of assessment. This is understandable, since it is addressing discipline-specific problems such as assessing understanding of probability or graphs. However, there are some important and useful ideas in the general assessment literature. We have already mentioned the notion that assessment shows students what we as teachers value, and hence directs their learning (as discussed by Ramsden, 1992, Wild *et al*., 1997, and Chance, 2002). While we do not disagree with this view, we believe that there are other, equally important, determinants of student learning, including their own interest in the material, and their perceptions of their future professional roles.

Assessing student learning in groups is an area that could receive more attention within the statistics education literature. Research located in the areas of psychology, education and management can give us insight into the pedagogical philosophies that underpin such an approach to learning and assessing. An analysis of the work of Jaques (2000), Johnson and Johnson (1994), Carr and Kemmis (1993), Salomon and Globerson (1989) and Vygotsky (1978) shows that group work and group assessment allows instructors to develop more comprehensive assignments; enables students to gain an insight into group dynamics and processes; allows students to develop interpersonal skills; allows students to be exposed to the viewpoints of other group members; encourages students to be prepared for the 'real viewpoint'; and promotes reflection and discussion as an essential part of the process of becoming competent and reflective practitioners. Moreover, group-work skills are viewed by employers as the most important generic attributes that students should develop to prepare for the world of work (Dearing, 1997). Wenger (2000) indicates that social activities are aligned with modes of *belonging*: this concept becomes important when students learn in groups as they are then more likely to be engaged with their studies, demonstrate creativity through an 'imagination' (p 226) and align their studies to their aspiration for work. However, the skills required for effective peer and group assessment (and indeed for self assessment) are not innate in students, and they should be discussed and practised in courses of study (Black, 2005) and supported by appropriate technology: a successful approach is discussed by McKenzie and Freeman (2002).

We return to our earlier point that any view of assessment, or indeed any specific assessment task itself, implies a particular view or theory of student learning on which it is based, usually implicit and only occasionally explicitly. So, for instance, Begg (1997) discusses the difference between assessment based on a behaviourist theory of learning and that based on a constructivist theory. He points out that the former implies breaking knowledge down into discrete objectives and testing whether students have mastered each, while the latter implies a focus on finding out what students have constructed for themselves, and using what processes. Budé (2006) bases his assessment of different levels of understanding of statistics on theories of cognitive psychology. Garfield and Ben-Zvi (2005) talk of an explicit constructivist learning theory with a 'spiral' approach to learning about variability, and point out that outcomes are best assessed in research studies using extended data investigations or open-ended survey questions: their suggested assessment questions provide alternatives more suited to standard classroom/lecture situations. As these examples indicate, it is important that the assessment used is coherent with the overall pedagogical approach and theory of learning on which it is based.

STUDENTS' CONCEPTIONS OF STATISTICS

In this section we will describe some results from studies that we have carried out by interviewing students about their conceptions of statistics and learning in statistics, and other studies that investigated students' conceptions of assessment. The views uncovered have clear implications for coherent assessment, which we will discuss in the following section.

We started with the aim of developing pedagogy that supports and develops students' learning of statistics from their own perspectives, and the most obvious way of doing this was to actually ask students how they understand statistics and how they go about learning statistics. We did this by carrying out several series of interviews with undergraduate students and recent graduates, majors in statistics and students who were studying statistics as part of another discipline. In these interviews, we asked students open-ended questions such as *"What do you understand statistics to be about?"*, *"How do you know when you have learned something in statistics?"* and *"What part do you think statistics will play in your future professional work?"* Students' responses were investigated with further probing questions. Altogether, we carried out interviews with over 80 students during 1999–2003 resulting in over 250,000 words of transcripts. The results of our investigations have been published in a series of papers (Petocz & Reid, 2001, 2003a, 2005; Reid & Petocz, 2002a, 2002b, 2003).

The theoretical basis for our approach was a methodology known as phenomenography: this looks at how people experience, understand and ascribe meaning to a specific situation or phenomenon (Marton & Booth, 1997; Bowden & Green, 2005). It is a qualitative orientation to research often used to describe the experience of learning and teaching, seen as a relation between the person and the situation that they are experiencing. Phenomenography defines aspects that are critically *different* within a group involved in the same situation, and its emphasis on (qualitative) variation parallels the emphasis that statistics itself places on (quantitative) variation. It is this variation that makes one way of seeing statistics qualitatively different from another, and allows definition of qualitatively different categories. The outcome of a phenomenographic study consists of the set of categories and the relationships between them: this is referred to as the *outcome space* for the research. Often, such categories show a hierarchical and inclusive relationship, in terms of the logical definition of the categories themselves and/or in terms of an empirical hierarchy. In the latter case, people who seem to hold the "broadest" conceptions also show an awareness of the "narrower" categories, while those who seem to hold the narrowest conceptions do not seem to be aware of any broader ones. This is, indeed, the reason why we, as educators, favour the broader, more inclusive categories over the narrower, more limited ones.

Here we summarise our findings: the details of our findings, together with data in the form of supporting quotations, are given in the papers referenced earlier. We identified six qualitatively distinct conceptions of statistics, which can be grouped into three levels from the most limiting (1) to the most expansive (6):

- *Focus on techniques:*(1) statistics is individual numerical activities, (2) statistics is using individual statistical techniques, (3) statistics is a collection of statistical techniques.
- *Focus on using data:* (4) statistics is the analysis and interpretation of data, (5) statistics is a way of understanding real life using different statistical models.
- *Focus on meaning:* (6) statistics is an inclusive tool used to make sense of the world and develop personal meanings.

Additionally, we identified six qualitatively distinct conceptions of learning in statistics, which can also be grouped into three levels, from the most limiting (A) to the most expansive (F):

- *Focus on techniques:* (A) learning in statistics is doing required activities in order to pass or do well in assessments or exams, (B) learning in statistics is collecting methods and information for later use.
- *Focus on subject:* (C) learning in statistics is about applying statistical methods in order to understand statistics, (D) learning in statistics is linking statistical theory and practice in order to understand statistics, (E) learning in statistics is using statistical concepts in order to understand areas beyond statistics.
- *Focus on student:* (F) learning in statistics is about using statistical concepts in order to change your views.

The outcome spaces that we have identified and described are empirically hierarchical and inclusive. Students who described the more limiting views of statistics or learning in statistics seemed unable to appreciate features of the more expansive views: however, students who described the more expansive views seemed to be aware of the narrower views, and were able to incorporate characteristics of the whole range of conceptions in their understanding of statistics and learning in statistics (transcripts of the interviews show this clearly). Further, these conceptions of learning in statistics seem compatible with results from other disciplines in higher education (e.g., Marton & Säljö, 1979, and examples in Marton & Booth, 1997).

In the area of creative arts education, Shreeve *et al.* (2004) have investigated students' conceptions of assessment, again using a phenomenographic approach. They showed their results in an outcome space with three hierarchical and inclusive levels. The narrowest level is ‘correction’, where students see assessment as a process done to them by tutors who check that they have done the right things and the right amount of work. A broader level is the ‘developmental’: here, students see the assessment activity as being designed to help their progress, with their tutors’ advice and help. The broadest level is ‘partnership’, where students see themselves as equal partners with their teachers in the process of evaluating and judging their own work. While this theoretical model was developed in the context of students of design and creative arts, it seems likely that these or similar conceptions would apply in statistics as well. However, the broadest ‘partnership’ conception doesn’t seem to be widespread in undergraduate assessment in statistics, though it certainly provides something to aim for!

University students generally look beyond their classes and curriculum towards their future professional life. Their perceptions of their future profession influence their approach to their learning at university, as indeed their lecturers’ perceptions of their professional world influence their teaching approach, and this link is important pedagogically. The idea of the *Professional Entity* (Reid & Petocz, 2004a) developed from a recognition that views of professional work and learning, and the relationship between them, had similarities across disparate disciplines – initially in music education (Reid, 1997), then design, law, mathematics and statistics – as well as some disciplinary variation.

The Professional Entity is a way of thinking about students’ (and teachers’) understanding of professional work using three levels of conceptions. The narrowest is the *Extrinsic Technical* level, in which people describe a perception that professional work consists of technical components that can be used when the work situation demands it. In statistics, this is shown by a view that statistical work is concerned with gathering statistical techniques for use in different situations. At the broader *Extrinsic Meaning* level, people hold that professional work is about developing the meaning inherent in discipline objects. In statistics, this is shown by the view that statistical work is focused on finding meaning in sets of data. The broadest view is the *Intrinsic Meaning* level, in which people perceive that their professional work is related to their own personal and professional being. In statistics, this corresponds to a view of statistical work as creating and modifying views of the world based on numerical evidence.

The Professional Entity is an important idea since each of its levels corresponds with a particular approach to the discipline and to learning (or teaching or assessment) in that discipline. For example, a limiting ‘technical’ view of the profession of statistician corresponds with a learning focus on development of atomistic and technical statistical skills – the ‘focus on techniques’ conceptions. By contrast, an expansive ‘personal’ view of the statistical profession enables students to focus their learning on the meaningfulness of statistics – the ‘focus on meaning/student’ conceptions. If students are encouraged to broaden their conception of statistics and the statistical profession, they will develop correspondingly broader approaches to learning (see Reid & Petocz, 2002a).

The latest interviews that we carried out were with students studying ‘service’ statistics as part of courses in engineering and sports science. The number of students studying statistics as part of another degree is much higher that those majoring in the discipline, so we felt that it was important to also investigate their views. The details of our analyses are given in Petocz and Reid (2005): however, our results indicate that students in professional disciplines that use statistics have essentially the same range of conceptions of statistics and learning in statistics as do students majoring in statistics. We were surprised by this finding, as we had expected differences, based on generally accepted wisdom that service statistics students are different from statistics majors.

IMPLICATIONS FOR ASSESSMENT

The phenomenographic approach implies a theory of learning that is focused on the variation in students’ conceptions of any particular phenomenon, and sees learning as the process of broadening views from narrower to broader conceptions. Indeed, the notion of ‘deep’ and ‘surface’ approaches to learning (essentially the same idea discussed in Garfield & Ben-Zvi, 2005) comes from early studies carried out by phenomenographers (e.g., Marton & Säljö, 1979). This view or theory of learning implies a particular approach to helping students to learn and to assessment. Individual assessment tasks, and indeed the whole assessment scheme, should be designed to display and acknowledge the range of variation, and to encourage students towards the broadest conceptions of the discipline and of learning.

Some practical principles for assessment follow immediately and these are discussed and exemplified below. First, in order to make students aware of the range of views of a particular topic, it is useful not only to explicitly acknowledge and discuss the fact that different students will have different views, but also to demonstrate this explicitly. We could ask students in a class or a lecture to write down a brief answer to a question, and then display the results in a form that the whole class can see. Using current technology, this can be done immediately, or answers can be collected and displayed at the next class. For example:

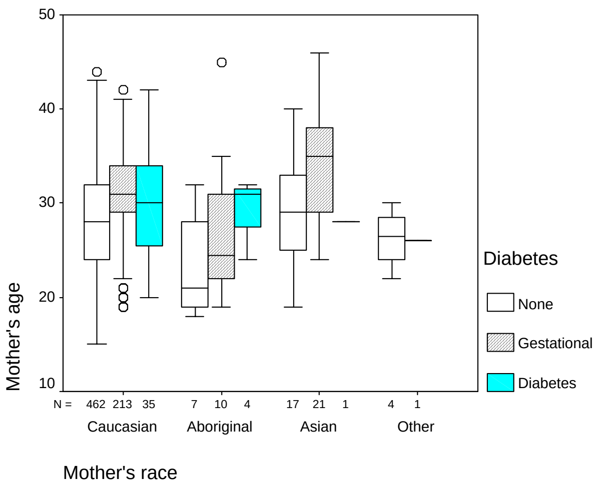
[1] In the film clip [from *Rosencrantz and Guildenstern are Dead*], they talk about 'examining their faith … at least in the law of probability'. What does the 'law of probability' say about coin tossing? Why do you think they got 79 heads in a row?" (see Bilgin & Petocz, 2006).

Further, encouraging students to work on assignment tasks in groups allows direct experience in variation of conceptions, and this is particularly useful in a context where this has already been discussed and experienced in classes. For example:

[2] *A question from a study of Hain et al. (1999), given in Wood and Petocz (2003).*
Discuss these questions in a group of three or four people. Do you think that this study provides evidence that Tai Chi has a positive effect on balance? On the basis of the results, would you recommend a course in Tai Chi to people that you know? Would such a course be particularly beneficial to older people? Incorporate the statistical evidence into your answers, but try to be clear about its limitations.

A further implication from an acknowledgement of variation is to give students opportunities to demonstrate what they can do, rather than asking closed-form questions. The next example comes from a first-year examination, based on a set of data about gestational diabetes:

[3] Look at the graph on the right, a clustered boxplot which looks at the relationship between mother's age, race and diabetes status. Write down *two* interesting features of the graph, and for each, explain why you think it is of interest.



Students will be encouraged towards the broader conceptions of statistics by assessment tasks that focus on analysis of data rather than specific techniques, and on the meaning of the analyses. This is not to say that techniques should be eschewed, but rather that they should be placed in the broadest personal or professional contexts, as in this example (Petocz, *et al*., 1996):

[4] In the visual quality test of Dow Corning's sealant, each sample is given a score from 1=Very Good to 8=Just Acceptable to 9=Just Unacceptable to 12=Worst possible. The results from a period of testing are shown below:

Visual quality test results Jan 94–Apr 95: 7 8 7 8 8 7 8 8 7 6 7 8 8 7 8 8 8 8 7 8 7 6 7 8 9 8 8 8 8 8 8 7 8 8 8 7 7 8 8 8 8 8 8 8 5 4 6 8 8 8 7 7 6 7 5 4 6 8 7 8 8 7 6 5 8 8 5 8 6 7 8 6 7 8 7 8 8 6 7 8 8 8 7 8 9 6 5 6 8 8 8 7 8 8 7 8 9 8 6 7 8 8 7 8 9 5 8 7 6 7 8 7 8 8 9 8 8 5 7 8 8 8 9 8 7 8 6 7 8 8 9 8 6 7 9 8 8 8 12 9 7 5 8 9 8 6 7 8 12 8 5 7 8 6 7 8 8 5 7 8 4 7 5 8 10 8 7 6 8 7 8 7

(a) Prepare a graph showing the frequency of occurrence of each value. (b) What does your graph show about the distribution of values? What conclusions could you draw? (c) Write a report in a few paragraphs to management. Explain the problem and suggest a solution. Give your reasoning in non-technical language that can be understood by people who don't know much statistics.

The ideas about Professional Entity that we described earlier imply that it is important to make explicit connection with students' future professions, and with professional skills such as communication and ethics. This is shown in the previous example, and also in the next ones.

[5] You have been asked to write a short piece for the university magazine. Write a short article of a few paragraphs (no more than 200 words long) describing some of the characteristics of the convicts of the First Fleet. Find an interesting title for your article. You may like to look for more information at http://www.gsat.edu.au/~markw/firstfleet/FirstFleetHomePage.htm.

[6] *A question from a study of Garavan (1997), given in Wood and Petocz (2003):*
The section describing the study procedures makes it clear that the receptionists in this study did not give their 'informed consent' for the study. Would you be happy to participate in a study of this type at your (present or future) workplace under these conditions? What aspects of the study would have to be changed to incorporate 'informed consent' of the participants?

Finally, in terms of broadening conceptions, self or peer assessment can be a very effective method. At the end of the projects discussed in Viskic and Petocz (2006), following discussion in class about the importance of such reflection for learning, students are asked to:

[7] Give a brief account of your investigations, describing the problems you faced and the successes you achieved.

Many quotes in our interviews show that some students are very aware of the conditions that support and encourage their own learning, and of the differences between surface and deep approaches to learning. This is shown in the following statement, given in Reid *et al*. (2005):

Julia: *There's monkey learning and there's proper learning. Monkey learning is finding out what you need to learn for the exam to get through, proper learning is finding out what's behind the numbers that you are writing down so that you know for yourself. There are people that do very well in a subject because they learn what they need to know for the exam, but you ask them three or four weeks later and they couldn't tell you. There are people that won't do that well in their marks, but you ask them three years down the track and they will be able to explain to you how that matrix works or whatever you are talking about. There's always a difference. And it takes a lot more time to learn the background than the 'what you need to know'.*

With such sophisticated insight into their own learning, it doesn't seem unreasonable to occasionally ask students to investigate aspects of statistics education themselves! For example:

[8] *From a study of Danaher (1998), a fairly mathematical article, from Wood and Petocz (2003):*
Work in a group of three or four people for this question. Some statisticians believe that statistics is a branch of mathematics, while others believe that it is a completely different discipline. Some university lecturers of statistics spend substantial time on the mathematical aspects of statistics, while others put in as little mathematics as possible. What do you think? Does your view depend on the type of statistics course, for example, a "servicing" course for another professional area, or a "major" course for students studying to be professional statisticians.

In terms of the whole assessment scheme, we would concur with many approaches that advocate less emphasis on timed examinations. Where university or department policy insists on such examinations, we would ensure that students are allowed to bring in a summary page of notes, so they don't have to spend their 'study' time memorising information. If the examination is required to have the multiple choice format, then we would agree with the ideas of Wild *et al*. (1997), for instance, to bury one false statement among several true ones:

[9] Which of the following factors would NOT influence the size of the sample that you would take for your study?
(a) The accuracy that you wanted for your results (for instance, estimating proportions to plus-or-minus 3% as opposed to plus-or-minus 5%).
(b) The number of subgroups that you wanted to study separately (for instance, males versus females; young, middle-aged or older people; low, medium or high income earners).
(c) The budget that you had for the experimental side of the project (for instance, money for interviewers to carry out the live interviews).
(d) The number of people in the population that you wanted to study (for instance, Sydney with 4 million people as opposed to Australia with 20 million people)

LOOKING FURTHER

Statistics education for future professionals can be usefully viewed as a higher-level numeracy – one of a range of key professional skills and dispositions such as communication, creativity, sustainability, ethics and cross-cultural sensitivity. Of course, this is true for statistics professionals as well as those in areas that make use of statistics. Presenting and assessing statistics at the broadest level gains synergies from such components of professional formation. Such an approach, illustrated in the examples in the previous section, allows students to develop and practise these professional skills and increases the relevance and interest of the course.

Tertiary students in a wide range of areas will need such skills in their professional life. Yet our studies of students' understanding of professional work and their conceptions of learning, summarised in the notion of Professional Entity (Reid & Petocz, 2004a), imply that students will only engage with those aspects of their university studies that seem to them to be relevant for their future careers. Without appropriate pedagogy, students may have trouble understanding the role of such professional attributes in their future professional life.

A statistics course – particularly a service course – presents an ideal opportunity to integrate these skills into their studies. In a sense, this is a broader version of Ramsden's (1992) principle that assessment drives learning. In such a statistics course, we have the opportunity to support students' professional formation by incorporating professional skills and dispositions into the curriculum and the assessment, and in return we benefit from students' increased engagement.

However, this is not without problems. Firstly, students may be unaware that creativity, sustainability, ethics or cross-cultural sensitivity (let alone communication) will be important in their studies or their professional life: for example, they may have imprecise ideas about creativity from assessment criteria that specify it as a high learning outcome (Cropley, 2001; Reid & Petocz, 2004b). Secondly, lecturers have a range of conceptions about these dispositions, some broad but others quite narrow (Reid & Petocz, 2006). Thirdly, it is not easy to design assessment tasks that focus students on these skills (ethics, for example, see Jebeile & Reid, 2002), or even incorporate them into the statistics curriculum (Petocz & Reid, 2003b).

It seems that the skills and dispositions that contribute to professional formation often appear to be something that 'can be taught in other subjects'. Even if lecturers discuss such issues, their students may perceive them to be peripheral topics until they develop more realistic notions about professional work. A statistics course aimed at the broadest levels of understanding, and with a correspondingly broad approach to assessment, can help students develop an understanding of the nature of their future professions, including the role of such professional attributes.

These attributes are aligned with the Intrinsic Meaning view of the Professional Entity and include the most expansive conceptions of statistics and learning as component parts. For instance, if we take just the idea of creativity, we need to consider how the term and activity is construed by teachers, students, and others in the profession. In the context of a first statistics course (for instance), is creativity the ability to solve, or find, a problem? Practically, how can creativity be assessed within a learning task if it has not been defined within that task? Is creativity the ability to see unusual applications for a set of data? Is creativity in statistics an issue that is discussed, or assessed, or even recognised? And how would you set up learning situations where this can occur? The development of these attributes enhances students' ability to find professional jobs and contribute in meaningful ways to the professional workplace. Such attributes are highly valued by the workplace, yet they are components of study that are often ignored or glossed over to allow more time for 'essential' components of content.

Using assessment to support learning, and in particular to help students develop their professional dispositions, is an important challenge for a statistics course – particularly a first, service statistics course. In the assessment examples we gave earlier, we have tried to indicate an approach, but we leave the challenge with you.

REFERENCES

Begg, A. (1997). Some emerging influences underpinning assessment in statistics. In J. Garfield & I. Gal (Eds.), *The Assessment Challenge in Statistics Education*, IOS Press, 17–25.

Bilgin, A. & Petocz, P. (2006). LAMS and statistics pedagogy. In R. Philip, A. Voerman & J. Dalziel (Eds.), *Proceedings of the First International LAMS Conference 2006: Designing the Future of Learning*, The LAMS Foundation, Sydney, 15–24, [Online] http://lamsfoundation.org/lams2006/papers.htm .

Black, P. (2005). Assessment for learning. Where is it now and where is it going? *Presented at Improving Student Learning through Assessment Conference*, London, September [summary Online] http://www.brookes.ac.uk/services/ocsld/isl/isl2005/keynote_black.ppt .

Bowden, J. & Green, P. (Eds.), (2005). *Doing Developmental Phenomenography*. RMIT University Press, Melbourne.

Budé, L. (2006). Assessing students' understanding of statistics. In A. Rossman & B. Chance (Eds.), *Proceedings of the Seventh International Conference on Teaching Statistics, ICOTS7*, Salvador, Brazil, IASE, [Online] http://www.stat.auckland.ac.nz/~iase/publications/17/6G3_BUDE.pdf .

Carr, W. & Kemmis, S. (1993) *Becoming Critical*. Falmer Press, London.

Chance, B. (2002). Components of statistical thinking and applications for instruction and assessment. *Journal of Statistics Education*, 10(3), [Online] http://www.amstat.org/publications/jse/v10n3/chance.html .

Cropley, A. (2001). *Creativity in Education and Learning.* Kogan Page, UK.

Curcio, F. & Artzt, A. (1997). Assessing students' statistical problem-solving behaviours in a small-group setting. In J. Garfield & I. Gal (Eds.), *The Assessment Challenge in Statistics Education*, IOS Press, 123–138.

Danaher, P. (1998). Estimating a cricketer's batting average using the product limit estimator. *The New Zealand Statistician*, 24(1), 2–5.

Dearing, R. (1997). *The Dearing Report – National Committee of Inquiry into Higher Education*. NCIHE Publications, Middlesex, UK.

delMas, R., Ooms, A., Garfield, J. & Chance, B. (2006). Assessing students' statistical reasoning. In A. Rossman & B. Chance (Eds.), *Proceedings of the Seventh International Conference on Teaching Statistics, ICOTS7*, Salvador, Brazil, IASE, [Online] http://www.stat.auckland.ac.nz/~iase/publications/17/6D3_DELM.pdf .

Gal, I. & Garfield, J. (Eds.) (1997). *The Assessment Challenge in Statistics Education*, IOS Press, [also Online] http://www.stat.auckland.ac.nz/~iase/publications/assessbk/ .

Garavan, T. (1997). Interpersonal skills training for quality service interactions. *Industrial and Commercial Training*, 29(3), 70–77.

Garfield, J. & Ben-Zvi, D. (2005). A framework for teaching and assessing reasoning about variability. *Statistics Education Research Journal*, 4(1), 92–99, [Online] http://www.stat.auckland.ac.nz/serj .

Garfield, J. & Gal, I. (1999). Assessment and statistics education: current challenges and directions. *International Statistical Review*, 67(1), 1–12.

Hain, T. C., Fuller, L., Weil, L. & Kotsias, J. (1999). Effects of T'ai Chi on balance. *Archives of Otolaryngology Head and Neck Surgery*, 125, 1191–95.

Jaques, D. (2000). *Learning in Groups*. Kogan Page, London.

Jebeile, S. & Reid, A. (2002). Constructive alignment in accounting education: evaluating the use of case studies and problem-based approaches in large classes. *Proceedings of the About Evaluations and Assessment Conference* (CD-ROM). Queensland Uni of Technology.

Johnson, D. W. & Johnson, R. T. (1994). Positive interdependence: key to effective cooperation. In R. E. Slavin (Ed.), *When and Why does Cooperative Learning Increase Achievement? Theoretical and Empirical Perspectives*. Prentice-Hall, Englewood Cliffs, NJ.

Lunsford, M., Rowell, G. & Goodson-Espy, T. (2006). Classroom research: assessment of student understanding of sampling distributions of means and the Central Limit Theorem in post-calculus probability and statistics classes. *Journal of Statistics Education*, 14(3), [Online] http://www.amstat.org/publications/jse/v14n3/lunsford.html .

MacGillivray, H. (2006). Using data, student experiences and collaboration in developing probabilistic reasoning at the introductory tertiary level. In A. Rossman & B. Chance (Eds.), *Proceedings of the Seventh International Conference on Teaching Statistics, ICOTS7*, Salvador, Brazil, IASE, [Online] http://www.stat.auckland.ac.nz/~iase/publications/17/6B4_MACG.pdf .

Marton, F. & Booth, S. (1997). *Learning and Awareness*. Lawrence Erlbaum, New Jersey.

Marton, F. & Säljö, R. (1979). On Qualitative Differences in Learning: I. Outcome and Process. *British Journal of Educational Psychology*, 46, 4–11.

McKenzie, J. & Freeman, M. (2002). SPARK, a confidential web-based template for self and peer assessment of student teamwork: benefits of evaluating across different subjects. *British Journal of Educational Technology,* 33(5), 551–569.

Pange, J. (2006). Assessing and educating preschool teachers on probability concepts in the classroom. In A. Rossman & B. Chance (Eds.), *Proceedings of the Seventh International Conference on Teaching Statistics, ICOTS7*, Salvador, Brazil, IASE, [Online] http://www.stat.auckland.ac.nz/~iase/publications/17/3H2_PANG.pdf .

Petocz, P., Griffiths, D. & Wright, P. (1996). *Statistics for Quality – Using Statistics in Australian Industry*. Video, 33 mins, booklet of exercises, Summer Hill Films, University of Technology, Sydney and University of Wollongong.

Petocz, P. & Reid, A. (2001). Students' experience of learning in statistics. *Quaestiones Mathematicae, Supplement 1*, 37–45.

Petocz, P. & Reid, A. (2003a). Relationships between students' experience of learning statistics and teaching statistics. *Statistics Education Research Journal*, 2(1), 39–53, [Online], http://www.stat.auckland.ac.nz/serj .

Petocz, P. & Reid, A. (2003b). What on earth is sustainability in mathematics? *New Zealand Journal of Mathematics*, 32 *Supplementary Issue*, 135–144.

Petocz, P. & Reid, A. (2005). Something strange and useless: service students' conceptions of statistics, learning statistics and using statistics in their future profession. *International Journal of Mathematical Education in Science and Technology*, 36(7), 789–800.

Ramsden, P. (1992). *Learning to Teach in Higher Education.* Routledge, London

Reid, A. (1997). The hierarchical nature of meaning in music and the understanding of teaching and learning. *Advancing International Perspectives*, 20, 626–31.

Reid, A. & Petocz, P. (2002a). Students' conceptions of statistics: a phenomenographic study. *Journal of Statistics Education*, 10(2), [Online], http://www.amstat.org/publications/jse/v10n2/reid.html .

Reid, A. & Petocz, P. (2002b). Learning about statistics and statistics learning. *Australian Association for Research in Education* 2002 *Conference Papers*. Compiled by P. Jeffery, AARE, Melbourne, [Online] http://www.aare.edu.au/02pap/rei02280.htm .

Reid, A. & Petocz, P. (2003). Completing the circle: researchers of practice in statistics education. *Mathematics Education Research Journal*, 15(3), 288–300.

Reid, A. & Petocz, P. (2004a). The Professional Entity: researching the relationship between students' conceptions of learning and their future profession. In C. Rust (Ed.), *Improving Student Learning: Theory, research and scholarship*, Oxford Brookes, 145–157.

Reid, A. & Petocz, P. (2004b). Learning domains and the process of creativity. *Australian Educational Researcher*, 31(2), 45–62, [Online] www.aare.edu.au/aer/online/40020d.pdf .

Reid, A. & Petocz, P. (2006). University lecturers' understanding of sustainability. *Higher Education*, 51(1), 105–123.

Reid, A., Wood, L., Smith, G. & Petocz, P. (2005). Intention, approach and outcome: university mathematics students' conceptions of learning mathematics. *International Journal of Science and Mathematics Education*, 3(4), 567–586.

Salomon, G. & Globerson, T. (1989). When teams do not function the way they ought to. In N. M. Webb (Ed.), Peer Interaction, Problem-solving and Cognition: Multidisciplinary perspectives, *International Journal of Educational Research*, 13, 89–99.

Shreeve, A., Baldwin, J. & Farraday, G. (2004). Variation in students' conceptions of assessment. In C. Rust (Ed.), *Improving Student Learning: Theory, research and scholarship*, Oxford Brookes, 223–234.

Sowey, E. (2006). Letting students understand why statistics is worth studying. In A. Rossman & B. Chance (Eds.), *Proceedings of the Seventh International Conference on Teaching Statistics, ICOTS7*, Salvador, Brazil, IASE, [Online] http://www.stat.auckland.ac.nz/~iase/publications/17/3A1_SOWE.pdf .

Viskic, D. & Petocz, P. (2006). Adult students' views of mathematics: reflections on projects. *Adults Learning Mathematics International Journal*, 1(2), 6–15.

Vygotsky, L. (1978). *Mind in Society: The development of higher psychology processes.* Harvard University Press, Cambridge, MA.

Watson, J. (1997). Assessing statistical thinking using the media. In J. Garfield & I. Gal (Eds.), *The Assessment Challenge in Statistics Education*, IOS Press, 107–121.

Wenger, E. (2000). Communities of practice and social learning systems. *Organization*, 7(2), 225–246.

Wild, C., Triggs, C. & Pfannkuch, M. (1997). Assessment on a budget: using traditional methods imaginatively. In J. Garfield & I. Gal (Eds.), *The Assessment Challenge in Statistics Education*, IOS Press, 205–220.

Wood, L. N. & Petocz, P. (2003). *Reading Statistics*. University of Technology, Sydney.