Improving Statistical Education through the Experience of Reflective Practice

A thesis submitted in fulfilment of the requirements for the award of the degree

DOCTOR OF PHILOSOPHY

from the

UNIVERSITY OF WOLLONGONG

by

Anne Louise Porter,
B Commerce (Hons App Psyc) NSW,
Dip Ed, Dip Comp Sci, Dip Math, MSc (Hons), Wollongong

School of Mathematics and Applied Statistics
and
Faculty of Education
2001

Declaration

In accordance with the regulations of the University of Wollongong, I hereby state that the work described herein is my own original work, except where due references are made, and has not been submitted for a degree at any other university

Anne Louise Porter

Abstract

This thesis is the recount of a study that began with the aim of unpacking the statistical expertise of the teacher and author, with the intent of improving statistics teaching and learning. In the process of doing this, the researcher examined the expertise of other experts through a case study of a statistics professor, concept mapping of ideas of statistics professionals and through an examination of statistical literature. As the researcher and teacher moved to a position of accepting that *statistics is a study of variation,* she discovered a failure by authors of many introductory textbooks to appropriately acknowledge variation as a (the?) fundamental statistical concept.

In the second phase of the research, the teacher constructed a pre-tertiary statistics curriculum and taught it to a cohort of 64 students. Drawing upon constructivist ideas, teaching was based on using activities to elicit statistical thinking in students. Exercises on measurement, sampling, probability, and questions about relationships were all used to illustrate the nature of variation. Within-session student assessment involved providing a written reflection upon the statistical ideas generated in class. The ideas needed to be exemplified in an everyday context. End of session examinations included a concept mapping exercise to explore students' understanding of major statistics concepts. The continued process of unpacking expertise and reflective practice led to the teacher modifying the curriculum for the second implementation. The theme adopted was *statistics is the study of variation as it occurs throughout the research process* (ethics, questions asked, design, measurement, sampling, description, analysis and drawing conclusions). For this cohort of 79 students, their understanding of statistical concepts improved, but their satisfaction with the subject declined. For the third implementation, the subject was again modified to include an explicit focus on *learning how to learn statistics*. In this third variation of the subject, the 61 students continued to have a good understanding of the statistical concepts but higher levels satisfaction with the subject.

During the various implementations, the teacher-researcher experienced shifts in interpretation of the educational literature. Students demonstrated in their reflective homeworks that the act of reflection was not automatic. Students' reflective homeworks often focussed on hard concepts or easy concepts, but not necessarily on the range of concepts. They often focussed on the particular or the general patterns but not both. When transferring to a new task, many retained a chronological ordering rather than a logical ordering of material. Some students perceived only one level of meaning, while others perceived multiple levels. Rather than perceiving a need to have students construct their knowledge (through activities), the emphasis shifted to selecting pedagogical techniques that would reveal what knowledge the students had constructed and how they experienced the learning environment. Students were now perceived as constructing knowledge no matter what pedagogical approach was used. The activity basis for teaching was retained, with the emphasis on the benefit of experiencing statistical thought through the completion of a task.

An explicit focus on learning how to learn was incorporated into the third implementation as students needed to be aware of the gains they had made in learning to learn. They needed to learn how to deal with negative emotions. Important and complicated ideas, easily acquired using the activity based approach, were often dismissed as too easy to be important, and the associated feeling was that they had not learned anything. Students needed to be able to handle the discomfort associated with uncertainty; uncertainty expressing ideas in writing; uncertainty associated with not being told what to learn; and uncertainty in not obtaining deterministic answers. Students needed to learn to write effectively. They needed to value the learning process and to recognise what they had overcome in the process of learning statistics, whether it was overcoming issues arising inside or from outside the classroom.

Based on the experience gained in this project, a model for improving statistical education in the broadest sense would involve a reflective practitioner methodology. Reflection upon student lives (current, past and future intent); curriculum (fundamental statistical ideas), the pedagogy, and how students experienced learning are all objects for reflection. Reflection also includes the comparison and contrast of experiences and understanding of the teacher with those reported in the literature. In this study, the key to improving students' understanding has been attributed to that part of the reflective process that has focussed on making explicit the fundamental concepts of the discipline. Improving students' affect has been attributed to the inclusion, within the activity-based pedagogy, of an explicit focus on learning how to learn.

Acknowledgments

I thank both my supervisors John Hedberg and David Griffiths. I have enjoyed this undertaking enormously. Each has contributed in their own special ways to this thesis. Their contribution has always been with good humour, and the team work, the criticism and encouragement has been greatly appreciated. A special thank to David Griffiths is warranted for being the object of my scrutiny for so many years, for opening yourself to criticism and review.

To the students of Gateway, thank you also for revealing so much of yourselves. Through you my eyes have been opened.

To my brother-in-law James Wallace, my sister Carol Williams and neighbour Sue Martin many thanks for your input, discussion and critical review. Your comments and all the other support you have provided me have been highly valued.

To my husband, Tony, thank you for seeing my studies as an eccentric form of fishing, for the cooking and for all the other ways you have contributed to this thesis and to our family life. Also to my children, Erica, Emily, Bruce and Christina, who have only ever known me as a student, thank you for thinking that a mum doing homework is the norm. To my mother who missed my company on many weekends, thank you.

Finally a thank you to Professor John Bain for it was a paper written by Professor Bain that inspired this thesis and which set me on a course of discovery that I may otherwise have missed.

Table of Contents

Section 1 An Introduction

Chapter1 A Journey Begins

Section 2 Unpacking and Repacking Statistical Knowledge

Chapter 2 Representing Knowledge Structures

Section 3 Implementing Curriculum

Chapter 4 Preparing to Teach

Chapter 5 Reflections in and on practice

Chapter 6 A Grounded Theory

References

List of Tables

Exhibits

Figures

Section 1

An Introduction

The teacher with his watering can
Of formulae and words,
Of theories, theorems, epigrams,
Hypotheses and surds,
With cunning, wisdom, wit and stealth
Pretends to rear the perfect man
Exactly like himself

But what he's really looking for
Are funnels stuck in ears
Down which he may hopefully pour
(to boost his ego, calm his fears)
his facts and figures, data, dates,
the knowledge he anticipates
will cancel all arrears.

"An Image from Montaigne" (Kellaway, 1922)

CHAPTER 1

A Journey Begins

> It takes two studies to present one in qualitative research. One is the official research project and the other study is the study of that study. (Chenail, 1995).

...and so it has been with this thesis. The final product is the culmination of a seven year study. The thesis began as a project initially titled *Unpacking statistical experts' knowledge in order to teach better.* It evolved into the thesis now titled *Improving statistical education through the experience of reflective practice*. The basis of the telling in this thesis is the story of the journey from one view of how to improve learning to another.

The story tells of a journey from perceptions and misperceptions of theories and methodologies to new perspectives. The research framework can now be readily described as a qualitative study (see for example Dey, 1993) involving a self study of practice (Schon, 1983) and as utilising a teacher-as-researcher (see for example Adler, 1993) mode of research. Reflection-in-action and reflection-on-action (Schon, 1983) are the primary mechanisms for making sense of the data and hence for developing a grounded theory as to how to improve statistical teaching and learning. The pedagogical approach stems from a constructivist philosophy of *coming to know*, intertwined with principles associated with experiential learning. The mode of inquiry can be justified:

> …by focusing on personal practice and experience, teachers may undertake genuine inquiry that leads to a better understanding of the complexities of teaching and learning. (Loughran & Northfield, 1996, p. ix).

This is not how the story began.

In one sense, the story had no definable beginning. Experiences as a student and teacher years before 'the study' helped shaped the work. The story also has an arbitrary ending, recognising that 'theory is an ever-developing entity, not as a perfect product' (Glaser & Strauss, 1967, p. 32). The ending represents a pause between further teaching experiences and growth in understanding, a pause in which this thesis was written. Furthermore, in using this narrative approach (see for example Clandinin & Connelly, 1991; Mattingly, 1991) to document the enquiry, it is *now* recognised that:

> Narrative inquiry is not the search for 'the' truth but, instead, a never-ending construction and reconsideration of situated meanings. While our encounters with reality may produce a meaningful and understandable flow of experience, what we experience is produced by the action of our organising schemes on the components of our involvment in the real world. (Diamond, 1997, p. 146).

The beginning of the project, is described by the formal aims and objectives, just as they were elaborated in the initial months of doctoral enrolment.

1.1 A Statement of Aims

The purpose of this study was to develop a grounded theory of the teaching of statistics. The intention was to explore the relationship between two themes. The first theme was that of 'unpacking the expert in order to teach the novice' better (Bain, 1990, p. 119). The expert of interest was the self as teacher although the method of making explicit this expertise involved examining the expertise of others. In his paper Bain, recognised that teachers need to understand or make explicit their own conceptions and skills in order to be able to teach better. The second aspect of the study involved elaborating those aspects of teaching and learning which characterise 'student-centred approaches', that is, understanding student conceptions and skills as they impact on their learning.

The following aims were proposed:

- to identify the fundamental concepts, the theory, and the logic that expert statisticians use in their construction of statistical knowledge;

- to unpack my own statistical knowledge, through the unpacking of other statistical experts;
- to develop curricula, directed at teaching novices, which have as their focus experts' key statistical concepts, theory and logic;
- to develop appropriate methodologies for the implementation of the curricula;
- to implement an introductory statistics curriculum; and,
- to develop a theory of teaching statistics which has been grounded both in the world of the statistical expert and in the reality of the classroom.

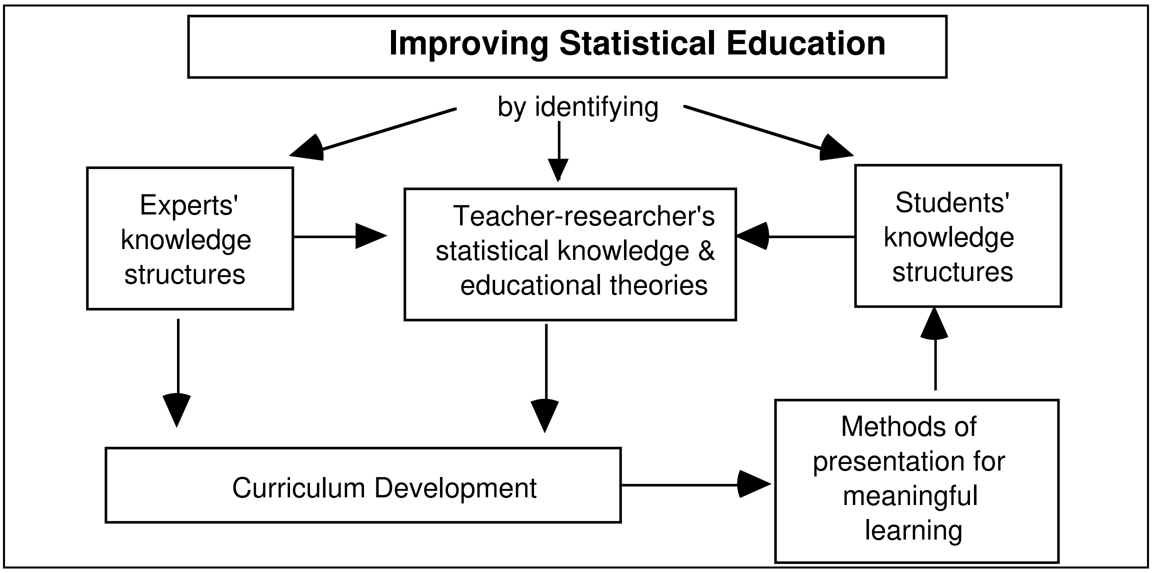
To fulfil the aims, the following questions were used to guide and frame the study:

- What fundamental statistical concepts guide experts?
- What aspects of statistical theory does an expert consider essential in the elementary practice of the discipline?
- What is the nature of the logic that statistical experts use?
- How are the logic, theory and these concepts combined?
- How does that logic, statistical theory and these fundamental statistical concepts differ between experts?
- How can these essential components of statistical experts' knowledge and knowledge processes be reflected in the statistics curriculum?
- What educational processes and structures need to be in place to assist learners of statistics?
- What student characteristics, both real and imagined, enable/inhibit student learning?
- What role does language, as widely defined, play in the learning process?
- How does the educational context influence the teaching and learning of statistical concepts?

1.2 An initial model for the conduct of research

A static representation of the model for exploring the improvement of statistical education through the unpacking of statistical knowledge is presented in Figure 1.1. The aim was to develop curriculum materials based on the central concepts identified by the unpacking of experts' and the teacher's statistical knowledge. Subsequently these were to be combined with appropriate theories as to how to educate. The curricula and appropriate pedagogical practices would define what and how materials were to be presented. If the statistical education was to be considered successful, the experts' conceptions could be expected to develop within the students' knowledge structures. Through the processes of unpacking expertise and developing and implementing the statistics curriculum, the operation of the model is dynamic, with feedback from students and experts leading to modifications of the curriculum.

Figure 1.1 The interplay between statistical knowledge, processes and student learning.



A continuous cycle of reflection, teaching and statistical practice, modifying teaching and more reflection ensured that the emergent theory of teaching statistics was well grounded in classroom practice.

1.3 Rationale for the Study

> Students frequently view statistics as the worst course taken in college. (Hogg, 1991, p. 324).

The reasons for this reaction to Statistics may be manifold. To counteract these reactions, many lecturers attempted to engage students by changing their pedagogical techniques. Strategies to enhance the learning of statistics have included using technology (Phillips & Jones, 1991; Stephenson, 1990; Yarbrough & Gilbert, 1999) and statistical learning packages (Marasinghe, Meeker, Cook, & Shin, 1996). They have included the designing of surveys (Winqvist, 1991), using real data (Carlson, 1999; Roberts, 1993), having students undertake projects (Green, 1991) and so use the data they have collected. Hogg (1991), for example, used projects. He says:

> …give students experience in asking questions, defining problems, formulating hypotheses and operational definitions, designing experiments and surveys, collecting data and dealing with measurement error, summarizing data, analysing data and communicating findings, and planning 'follow-up' experiments suggested by the findings. (p. 342).

Students, particularly those from the discipline of Psychology, who sought help in the University of Wollongong Learning Development Centre, often had their statistics learning supported by laboratory exercises and projects in which they designed and carried out experiments. These students recognised that they could calculate and write up statistical tests without ever being sure what it all meant. As one student who sought help in our University Learning Development Centre stated:

> Why I am here is to learn how to analyse questions and get a better idea of the concepts behind stats - I can follow formula but have no idea of the concepts. (Porter, Griffiths, & Hedberg, 1993).

A second Psychology student, Mary, sought help despite getting 100 % in her statistics tests. The test included questions requiring calculation of the standard deviation, independent t-test (t-test for use with independent samples) and related t-tests (paired t-test). Mary said:

> I got lost when the big letters eg 'X' and the little letters 'x' were used. I got through because I can extend from mock up tests to tests but I did not understand the whys and wherefores, and at times the tutor/lecturer would write one thing and then say 'wrong one' e.g. 'X', 'x' scrub it out and write the other. Standard deviation, standard error, what is the difference? . (Porter et al., 1993).

Difficulties in engendering meaningful learning are evident across the educational spectrum. Novak and Gowin (1984) described this lack of meaningful learning in a typical educational situation. Students engaged in methodological or procedural activities were characterised as follows:

> [They] are usually not consciously guided by the kinds of conceptual and theoretical ideas scientists use in their enquiries - there is no active interplay between the thinking… and the doing… (p. 56).

Evidence of this failure to engender meaningful, as distinct from rote learning, can be found in the many disciplines in which Statistics is taught, whether it be Psychology, Economics, Geography, Sociology, Biology, Education and equally in the Statistics discipline itself and elsewhere.

It is the perception of many statistics educators that engendering meaningful learning in the student of Statistics is not readily accomplished. Students can often master the different statistical procedures but lack understanding of the theory underlying these procedures. (See for example Karake, 1990; Wells, Pollatsek, & Boyce, 1990). As Burghes (1993) said of his students:

> I can get my students to pass their exams ... but so far I'm sure that I haven't been able to give them any real understanding and feel for the subject. (p. 68).

The principal mechanism that was employed to assist students seeking help in the Learning Development Centre involved the examination of Statistics through frameworks (see Chapter 2 for examples). Brew and Boud (1995) characterised this approach to teaching as one which:

> …stresses the need for students to organise their ideas into a framework. The students' task is to deepen their understanding by refining their sense of the way the ideas link together in an overall pattern. The emphasis is on understanding rather than just on memorising and application. (p. 266).

This approach to remediation was an attempt to counteract the pervasive sense that students sometimes have of statistics being an endless collection of rules and formulae, with little connection between them. Corresponding to this, is the conception of knowledge identified by Brew and Wright (1990). In this study, lecturers and tutors were described as seeing:

> …knowledge in terms of a set of more or less discrete facts which were then applied. The students' task was seen as learning the facts in order to use them to solve problems. The implication was that students have to learn to assimilate the facts and also learn to apply them… The assumption is that the more facts one has the better one is able to solve problems. Not only that but the facts tend to be viewed as separate. (p. 206).

Frameworks helped elucidate the content or process the students were required to learn. However, the remedial teacher continually questioned 'Why are students being asked to do this?' When moving from an academic role to that of a statistician, Pike (1994) asked, 'Why is the reality so different from the expectation?' There appeared to be a mismatch between the questions and procedures undertaken by students in courses and those undertaken through the practice of statistics or statistical consulting. This observed discrepancy was one of the driving forces for the orientation toward unpacking experts' knowledge in order to appropriately structure ideas as they were introduced to students.

1.4 Stages of research

The stages of the study and methodological approaches are summarized in Table 1.1. The table represents what emerged as the design, or course of action, rather than what was planned. The initial phases of research included:

- identification of the methodological approaches or frameworks for undertaking the research;
- positioning of myself as researcher in relation to the research; and,
- finding ways to represent knowledge structures and the unpacking of statistical expertise in order to identify central statistical concepts (an ongoing pursuit).

These initial phases were followed over a period of 18 months by three sequential implementations of statistical literacy curricula, providing the primary source of data for educational theory building. Over the subsequent two years, there was a search for a unifying theoretical and methodological framework that could encompass what had transpired. During this time, the practice of teaching in the same and varied contexts allowed for verification and extension of the emergent theory. The final phase involved finding a unifying framework for the findings. This required a sustained period of reflection on what had transpired throughout the study; reflections as to the nature of reflection itself and, finally, identification of areas of accord and discord between ideas generated by the grounded theory approach and that of the theories of others.

Table 1.1: An emergent design

Phase 1: Initial Research frameworks
- Qualitative research, grounded research, ethnography

Phase 2: Positioning myself as a teacher, learner & statistical consultant

Phase 3: Unpacking central statistical concepts
- Finding ways to represent expertise
- Unpacking of expertise: self, professor-expert and other experts
- Using the tools observation, discussion, questioning, concept mapping, Gowin's V heuristic, developing frames, mapping text, joint writing, multimedia development and literature analysis

Phase 4: Implementation: theory development

Cycle 1 Baseline implementation: Initial curriculum
- Develop pedagogical practices: experiential, activity based, no text, language immersion, silence, small group, videos & worksheets
- Implement, reflect, assess, observe and ask students, action research

Cycle 2 Curriculum more coherent : 'Statistics is about variability'
- Implement, reflect, assess, observe and ask students, action research

Cycle 3 Focus on learning how to learn
- Implement, reflect, assess, observe and ask students, action research

Phase 5: Teaching in other contexts: verification and extension of theory
- Statistical literacy for law students
- Statistical literacy for arts students.
- Introductory statistics subjects

Phase 6: Finding a unifying framework for improving the learning of Statistics
- Reflection on Reflection: addresses 'what was the mechanism of change?'
- Data accord and discord with other theoretical positions

1.4.1 Initial research frameworks: representing and unpacking expertise

When the thesis began, the researcher was captivated by Bain's (1990) suggestion that, in order to improve our own teaching, teachers' should make explicit their own implicit understanding of their discipline. This study was conceived of as the ultimate irony: a qualitative approach (see for example Dey, 1993) was chosen in order to study a quantitative discipline. The researcher, a psychologist re-educated as a statistician and steeped heavily in experimental methodology, chose to explore non-experimental ways of coming to know. The act of research was an exploration of a different form of knowledge building. The 'primacy of emphasis' in this research was to be on the construction of theory rather than on the verification of theories and hence on qualitative approaches rather than quantitative methods (Glaser & Strauss, 1967).

In the first phase of this study the teacher-researcher chose to unpack her own expertise primarily through the exercise of unpacking, in depth, another expert's knowledge, a Professor of Statistics (and Rhodes scholar). The exercise remains necessarily incomplete. Both the expert and novice continue to develop their knowledge structures, as does the discipline itself. Indeed, the aim of unpacking an expert's expertise was, in one sense, foolish. Pick an expert, many years senior in statistical experience and unpack what they understand and represent it in some simple manner! One of the first books read on unpacking expertise, with a title long since forgotten, was based on such a study with the researcher stalking the object of study in the wee hours of the morning, watching them work in their lab. This is bizarre, to say the least. In this instance, the method of unpacking expertise was pervaded by a sense of controlled chaos. As is often characteristic in grounded research, the method was at times formal, meeting at a set time and place with a technique such as concept mapping. At other times, it was the recall of interactions that took place many years earlier, observation, or simply by being able to work on joint research and teaching projects with the expert. At times a diary was rigorously kept; at other times it lapsed.

Armed with an arsenal of tools for representing (described in Chapter 2) and unpacking expertise (described in Chapter 3), recognising the vastness of the unpacking necessary, the approach embodied what might be called a lifestyle approach to research. It could at times be considered to be an ethnographic approach.

The unpacking has been of sufficient depth to encompass designing the curriculum of both a statistical literacy subject and an introductory tertiary level statistics subject. In accord with grounded theory the data collection can stop when new conceptualisations emerge (Creswell, 1998). There has been a pause in the unpacking to permit writing.

The second descriptor of the research methodology was that of grounded research (see for example, Creswell, 1998; especially Glaser & Strauss, 1967; Miller, 1995). Rather than setting out to verify a particular theoretical position, the teacher-researcher was to immerse herself in data and, from it, develop theories about what transpired. The aim was to develop a theory of teaching statistics that was grounded both in the world of the statistical expert and in the reality of the classroom. In this manner, the theory should 'fit the task and work' (Glaser & Strauss, 1967), and hence be a useful guide to improving teaching and learning in the Statistics classroom.

The study was to have two phases: identifying statistical experts' knowledge to build a curriculum; and secondly, implementation of the curriculum and evaluation. It was the continued iteration between these two phases, examining experts' knowledge, and seeking answers and evidence as to how best to teach that a theory of how to teach statistics was to emerge. Fieldwork, observations, documents and a number of other strategies were to be used as a source of data from which the theory could be derived. With the experience of implementing the curriculum (refer Chapter 5), the emphasis of the thesis shifted from unpacking expertise to a focus on how students learn. In accounting for the changes, the emphasis of the thesis again shifted as the teacher-as-researcher examined the reflective process through which change was engendered (Chapter 6).

In undertaking research in this manner, there was a sense that prior theoretical knowledge could impede the process of developing grounded theory. The experience of moving from psychologist to statistician had involving relearning statistical language and many statistical concepts. Therefore, in the early stages there was no attempt to further develop a theoretical framework nor methodological approaches, although there was an attempt to clarify existing thoughts. However, consistent with the grounded theory approach, there was a perceived need for the researcher to set aside theoretical ideas in order that a substantive theory could emerge (Creswell, 1998) and hence there was early recognition of the need to position myself in relation to the research.

1.4.2 Positioning myself

'Instead of doing and writing up 'a' piece of qualitative research the 'I' becomes part of the research' (Geertz, 1983). As Guba (1990) suggested, one's paradigmatic groundings both shape and conceptually limit the research endeavour. For this reason, I, the teacher and researcher, initially made an extensive statement regarding the nature of acting as a good or bad teacher, learner, and statistical consultant, in order to position myself in relation to the research. This was an attempt to set aside my beliefs about statistical expertise, teaching and learning. It was an attempt to free myself from being theory-bound, so that I could identify the expert's knowledge structures in contrast to my own. Further it was so that I could observe what took place in the classroom in order to develop from the data a theory as to how to improve the learning of statistics.

In the reflection phase of the research, these initial positioning statements were to provide a source of data. These data could be examined during the reflective phase to see whether or not expectations were confirmed or rejected.

1.4.3 Literature search: purpose

In grounded research, drawing upon observations of the data is pre-eminent in developing theory. In this regard, one needs to be careful in drawing upon literature and the theories of others at too early a stage. Marshall (1990) requires that in a good piece of qualitative research:

> … it is clear that there was a phase of 'first days in the field' in which a problem focus was generated from observation, not from library research. In other words, it is a study that is an exploration, not merely a study to find contextual data to verify old theories. (p. 194).

In this thesis the initial purpose of the literature search was to:

- guide the process of unpacking expertise, by providing controversial issues which could highlight the logic used when an expert chose one approach in preference to another; and,

- identify contradictions, controversies, alternate approaches and problems in current statistical thought and statistical education practice.

The review was also very much a part of grounding of the emerging theory of teaching statistics. When issues emerged throughout the process of unpacking expertise, readings would centre upon those issues. In a similar manner, at the end of each implementation of the curriculum, the literature was searched for articles which 'spoke to me'. These articles seemed to convey a feasible explanation for what had been observed in class. In the next session's teaching, the curriculum and teaching methodology were modified as the reading, reflection and further consultation with the expert suggested.

Further, the literature was used to identify 'the comparisons that the author has forgotten or "thrown away" because of his initial focus' (Glaser & Strauss, 1967, p. 90). Thus, in the final phase of the research, it was possible to reflect upon areas of accord and discord between emergent themes, preconceptions and those suggested by the literature.

1.4.4 Identifying and representing central statistical concepts

Prior to the first implementation of the curriculum, there was an intensive period of searching for ways to represent knowledge and techniques to assist in the unpacking of statistical expertise. This activity became intertwined with the implementation phases of the study as teaching itself was opportunistic, commencing before the unpacking phase was complete. Further, observations of the teaching and student learning outcomes and processes inspired further unpacking of statistical expertise.

1.4.5 Implementation of curriculum

The first implementation of the *Mathematics and Statistical Literacy* subject[1] provided a base level of performance, but initiated a shift in the curriculum so that it was more coherent and closer to one expert view that statistics is a study of variability. The second implementation of the subject used a modified curriculum. Throughout these first two implementations, questions which emerged and re-emerged were 'how it is we come to learn?' and 'what do theories or philosophies have to say about the construction of knowledge?' The third implementation of the subject, by then titled *Learning Mathematics and Statistics*, had the students focus more heavily upon how it is that they, as students, learned. However, the requirements for the mastery of the statistical and mathematical materials remained the same as in previous implementations.

1.4.6 Teaching in other contexts: verification and extension of theory

The three implementations provided the majority of the formal data for this thesis. However, implementing modifications of 'the curriculum' continued. Ongoing teaching to other classes stimulated reflection upon the data. The change in teaching context provided an opportunity to examine specific aspects of teaching and learning statistics. The teaching in this phase involved:

- *Learning Mathematics and Statistics* for more groups of tertiary preparation students with a specific focus on students' interpretations of feedback;
- Statistical literacy for arts students;
- Statistical literacy for law students; and,
- other introductory statistics programs (informatics - mathematics and information technology students, for engineers, and for scientists).

---

[1]The reference to mathematics is often omitted as the focus of the work is on the statistical component of the course.

These additional implementations continued to inform practice. They inspired reflection upon the three detailed case study implementations and allowed the possibility of verification and extension of emergent themes.

1.4.7 Finding the unifying research framework: reflection on practice

The final phase, extending during and beyond the period where implementation continued to inform practice, involved a period of reflection on practice or more broadly reflection upon what had transpired throughout the study and subsequent teaching. A journal entry 7/12/94 reminded me that in terms of the basic statistical concepts and the emergent theory of teaching and learning, the outcomes were the ends sought in this study, not a story about the methodology. However, the process of reflection, the action of being teacher-researcher, reflective practitioner or undertaking a self study, became more than additional emergent descriptors of the methodological approach. They became than the means of connecting and directing the data gathering and interpretation throughout the study. Reflection during the practice of teaching or *reflection in practice*, as coined by (Schon, 1983), was identified as one of the major mechanisms for improving learning in the classroom. This was seen as a major shift in focus, and thus the story about the study became enmeshed with the study itself.

When reflective practice was coined as the third methodology descriptor it was loosely described as thinking about what had transpired in the classroom. Around this time a colleague commented, 'I think about it, then I change my teaching. But it doesn't make any difference'. Thus, the final stage of the research was spawned. As other researchers have asked (see for example Bengtsson, 1995; Eraut, 1995; Korthagen & Wubbels, 1995; Lucas, 1996; Naysmith & Palma, 1998; Quicke, 1996; Van Manen, 1995), so it was asked of this research 'What is nature of reflection and reflective practice?' The question became 'What is it that I or the students have done that has led to an improvement in learning outcomes?' 'How can this process be communicated to teachers who may wish to change their teaching and improve it?' This phase has involved an up-ending of the reflection process, spawning questions in this thesis such as: how has this reflection taken place?…what triggered it?… and, was it the practice of teaching, remnants of theory, or philosophical positioning or some other aspect that directed the reflection? What do other researchers have to say about the nature of reflective practice? This phase is detailed in the final chapter.

1.5 Philosophical assumptions

During the phase, *positioning of the teacher* (as described in Chapter 4), there was an oscillation between what was interpreted as different theoretical or philosophical positions regarding teaching and learning. There was no satisfactory resolution at the time, nor indeed was there an attempt to delve more deeply into the teacher-researcher's own theoretical or philosophical position by researching, then adopting, one or other position. Experience over fifteen years in the fields of psychology and education formed the basis of how to proceed: a more recent focus in statistics had allowed the distinctions between educational and psychological theories to recede. To adopt a particular theory would have run contrary to the aim of developing theory from the data and, as Glaser and Strauss (1967) suggest, would have led to the risk of forcing the data. Rather, there was an attempt to describe beliefs and to leave the distinctions between the theories a little blurred in order to see what emerged. That which emerged could then be viewed as different or similar to the initial position.

However, in general, the philosophical leanings were, and remain, in accord with a constructivist view of the how knowledge is constructed. Initially that view was not greatly deeper than the claims made by radical constructivists:

> 1. Knowledge is actively constructed, not passively received, either through the senses or by means of communication.
>
> 2. Coming to know is an adaptive process, tending towards fit or viability. It serves to organise one's experiental world, not to discover an objective reality. (von Glasfield cited in Barnes, 1994, p.1).

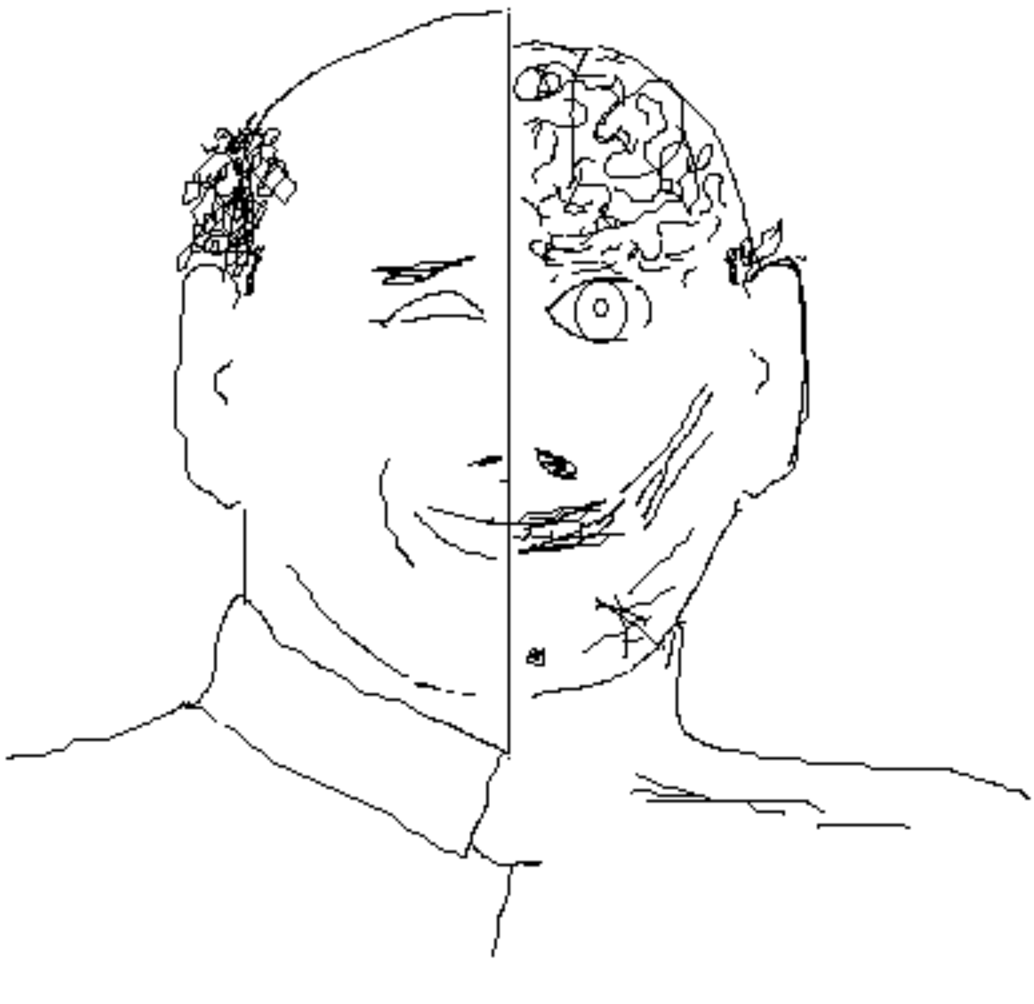
Associated with this view and integral to this study was the view of the role played by reflection. As Barnes (1994) said, 'engaging in activity of itself does not bring about learning'. It is reflection that highlights the conflict between the old knowledge structure and the demands of new information to be accommodated within it. These ideas were initially interpreted, as others have, as implying a particular mode of teaching, namely, a constructivist mode of teaching, wherein students were to construct their own knowledge and to reflect upon it. From this accident of interpretation, which is discussed in Chapter 5, flowed insight into the students' learning, a questioning as to the ontological and epistemological assumptions regarding the nature of coming to know. These were questions such as those asked by Stevens (1994) when examining implications of the various philosophical assumptions underlying different forms of constructivism. Stevens asked 'Does an external reality exist?', 'Does it exist independently of the observer?' and, 'Is it possible to know this reality objectively?' In the context of this work another question was of concern; 'What were the implications of these different philosophical positions for the unpacking of expertise and for student learning?' Throughout the study terms glibly adopted and assumptions held became challenged. As has already been alluded to, 'What does it mean *to reflect*?' and, 'What role does *reflection* play in learning?' became closely scrutinised questions, as they have for others. At this point, the sources of ideas and interpretations of the literature became an issue, as the teacher-researcher found many nuances in what others had written. Through the practice of teaching, theories about how students learned were beginning to emerge or be rewritten.

1.6 The next phase of the journey

Chapter 2 documents an early phase of the research; the search for tools to explore and represent knowledge. The aim was to establish a mechanism for the description of the central concepts and the theory and the logic that expert statisticians use in their construction of statistical knowledge. However it was soon recognised that there was a link between mechanisms for representing knowledge and theories and models of 'coming to know' or learning.

Section 2

Unpacking and Repacking Statistical Knowledge



(Writings, October,1993)

CHAPTER 2

Representing Knowledge Structures

2.1 Introduction

Through this thesis, I aimed to improve learning by identifying central statistical concepts and building these into the curriculum; to implement that curriculum; and, flowing from this, to develop a theory of statistical education. There were three sources of statistical expertise to be mined. The initial orientation of the study was to unpack in depth the expertise of one statistical expert (henceforth called the professor-expert). This was supplemented by a more general exploration of the expertise of others (referred to as other experts). However, in accord with Bain's theme, the ultimate aim was to unpack my own statistical expertise in order to improve teaching and learning in my classroom.

From the outset, it was apparent that the true structure of any expert's knowledge (even my own) is not readily accessible, nor is it packaged in a visual form to be instantly articulated. In the initial months of the research project, the emphasis was on identifying and use of tools to help elicit and represent statistical knowledge. Associated with this was the need to establish a methodology that resulted in valid representations of the experts' knowledge.

Given cognitive scientists' attempts to answer questions about the nature of knowledge, its components, its development and its use (Gardner, 1985) or on *coming to know* (West, Farmer, & Wolff, 1991) it is not surprising that many of the ideas and techniques developed by them were useful in eliciting and describing knowledge structures. The tools and techniques for eliciting, representing and utilizing knowledge were drawn primarily from the three areas:

- knowledge aquisition for expert systems, where 'the goal of knowledge aquisition is to model the knowledge of one or more experts in a way that will allow it to be encoded into an expert system.' (Brule & Blout, 1989, p. 1);
- psychology, especially cognitive psychology (see for example, Anderson, 1985; Eysenck & Keane, 1990; Stillings et al., 1995), in which tools have been developed to elicit reasoning that is often inaccessible, unconscious or intuitive (see for example, Brule & Blout, 1989; Fransella & Banister, 1977); and,
- instructional design, where the instructional process is to develop students' knowledge, including learning how to learn, developing students' awareness of appropriate cognitive strategies, and the identification of appropriate cognitive strategies for use in the 'teaching-learning transaction' (West et al., 1991).

Whilst these tools were used initially to explicate and describe others' expertise, it was anticipated that they would later be used in both the teaching and learning of statistics. Indeed, the teacher-researcher's personal experience of successfully helping students understand statistics involved jointly developing frameworks which could encompass the statistical ideas and procedures they were learning. Hence, the techniques used to represent statistical knowledge, and from this to develop the statistics curriculum, were also influential in both the implementation and evaluation of student learning. In this manner, an implicit connection was made between the tools selected for the purposes of explicating and representing expertise, and theories of how it is we *come to know*. As West comments, 'the cognitive strategies are a collection of known ways that people learn' (p. 26). This is problematic in relation to the purpose of developing a theory of statistical education grounded in the data of the classroom. The question remained 'Would the emerging theory of improving statistical education be attributable to the methods used?', that is, 'Would the theory be methodologically bound?'

In this context, the purpose of this chapter is to:

- describe more of the story that was thus begun in the first year of research;
- describe the sources of expertise which were to be unpacked;
- establish a methodology which would provide plausible representations of knowledge;
- examine the initial positioning of the teacher as researcher in relation to the theories of how we *come to know*;
- identify different types of knowledge;
- identify the techniques and strategies which formed the kit for unpacking and representing different types of knowledge; and,
- indicate how some of these seemingly simple ideas took on new meanings throughout the course of this thesis.

2.2 Sources of expertise for the unpacking

The levels of statistical knowledge of the experts to be studied were not measurable. However the experience of the professor-expert, other experts and the participant researcher can be characterised as follows.

The professor-expert

The professor-expert whose knowledge formed the focus of an in depth case study was a former Rhodes scholar and statistician with the Commonwealth Scientific and Industrial Research Organisation (CSIRO), a Professor of Statistics, with experience in both teaching and consulting in Statistics, and with a number of diverse scholarly publications.

Other experts

The exploration of *other experts* supplemented and provided a contrast to the professor-expert's knowledge. The access to these experts was, to some extent, opportunistic. Their qualifications and work experiences could not be formally assessed. The other experts were those who in some way proffered their statistical know-how. Contact with these experts was typically through discussions of this study with visiting scholars or through interaction at conferences. These discussions were often illuminating and unpredictable in outcome. Debates in the literature also provided sources of expertise, as did textbooks. Formal exercises conducted with statistics scholars at conferences and seminars also added to the understanding of possible structures for statistical knowledge although the experts were not separately identified.

The participant researcher

In accord with Bain's theme the ultimate purpose was to unpack my own statistical expertise. This was ultimately linked to the process of unpacking the knowledge of the professor-expert and the other experts. At the time that this project began, the participant researcher had completed a Bachelor of Commerce with Honours in Applied Psychology, Diploma in Education, Diploma in Computing Science, Graduate Diploma in Mathematics and an Honours Masters of Science. These latter qualifications both involved a specialisation in Statistics, whilst the first two involved completing service based subjects in statistics. At the commencement of this work, the participant-researcher held an appointment as a lecturer in a University Learning Development Centre assisting graduate and undergraduate students, primarily in the field of Statistics. In the two years immediately prior to commencing this thesis the researcher's primary role was as a statistical consultant for academics and graduate students in a Faculty of Health and Behavioural Sciences. For the previous ten years, multiple roles included those of teaching, research, administration, computing, and statistical consulting.

2.3 Creating an unpacking environment

The approach taken to unpack statistical expertise recognised that the unpacking of expertise would be an intensive process. The selection of the professor-expert in the first instance involved establishing a supervisor/student relationship with someone whose teaching and consulting had been observed for a period of five years. During that observation, the professor-expert had revealed himself as someone capable of illuminating statistical theory and practice. Not only could these previous interactions be drawn upon, but a commitment to interaction with the researcher was made for an an additional period of six years.

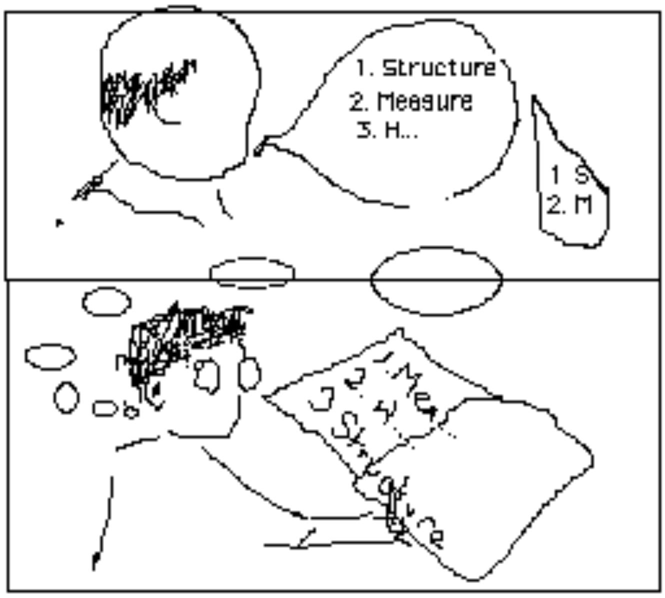
From the outset, a collegial environment was intentionally established. This permitted the opportunities for unpacking expertise to extend from formal data gathering exercises to that of naturalistic settings (shared research projects, teaching support, joint publications and committee work) utilising ethnographic methodology.

Whilst the case study began as a study of other it was quickly evident that the study of self was inextricably bound to the examination of the other. The search for a coherent statement of fundamental concepts and their linkages was driven by the participant-researcher making sense (and ultimately her students making sense) of this expert's knowledge and logic. It was in this manner, that the 'I' was afforded the role of participant researcher in the study rather than that of neutral observer. Thus, the methodology for unpacking expertise necessarily drew upon a participant-researcher model of research. In the words of Dey (1993):

> Analysis often proceeds in tandem with data collection, rather than commencing on its completion. The resulting analysis is contingent in character, since it in turn stimulates and is modified by the collection and investigation of further data. The researcher meanwhile becomes a participant in his or her own research project, for their own interpretations and actions become a legitimate object of subsequent analysis. Information on the researcher's own behaviour and thinking, in the form of fieldnotes, memos, diary or whatever, can become a vital source of data for the analysis. (p. 37).

However this method of research has to contend with the possibility of bias due to the researcher's participant role. As Dey discussed, qualitative researchers may use observational methods which produce data inconsistent with how subjects experience, perceive or explain events. Subject's explanations of events may involve unrecognised assumptions. And so it was experienced in this study, as captured by the cartoon in Figure 2.1. After a session with the professor-expert the researcher, acting as a scribe, made a simple translation in ordering two concepts. This was seemingly a slight change, but for the professor-expert it represented a significant change in meaning.

Figure 2.1 Changing meanings in the unpacking phase. (Writings, 1993)



2.4 Validity

When collecting and analysing qualitative data, safeguards are necessary to protect against fabricating evidence; discounting evidence and misinterpreting evidence. All empirical studies must address the challenge 'how do they know what they claim to know' (Schon, 1991, p. 343). For it is important to note that the neutral observer is a myth. External validation is often not possible although corroboration by other collaborators is possible.

> The trouble is that we tend to see what we want to see, and hear what we want to hear...Miles and Huberman describe 'plausibility' as the 'opiate' of the intellectual. (Dey, 1993, p. 222).

Given the interpretative nature of one researcher explicating another's knowledge structure, a procedure that involved the expert in commentary upon the unpacking of expertise was deemed necessary. The established collegial environment provided many opportunities for interaction and the corroboration of interpretations made of the professor-expert's understandings of statistics.

Throughout the unpacking phase, the professor-expert often provided a commentary upon the researcher's interpretations. This at times involved the professor-expert in altering concept maps or frames which were based on his draft statistics textbook but constructed by the researcher. At other times the commentary came through jointly writing scholarly papers (Porter, Griffiths, & Hedberg, 1993; Porter, Griffiths, & Hedberg, 1994a; Porter, Griffiths, & Hedberg, 1994b; Porter, Griffiths, & Hedberg, 1994c) or through critiquing papers written by the researcher (Porter, 1993a; Porter, 1995; Porter, 1996a; Porter, 1996b). The collaborative process established to unpack expertise became an importance source of corroboration of interpretations. The meaning that was attributed to the professor-expert's actions, words and omissions in publications were vetted and modified where necessary. This cyclical loop involving the professor-expert and researcher is one of the characteristics of an action research paradigm and necessary to ensure a reasonable representation of the professor-expert's views.

The process of unpacking the professor's statistical expertise commenced prior to the development and implementation of a statistics curriculum. However, it continued later in tandem with teaching and observation of the students' learning processes. The focus was sometimes the expertise of the expert professor, sometimes the less detailed representations by many other experts, and at other times the students' conceptions and misconceptions.

It is highly likely that most experts themselves have not come to closure on certain topics. It was apparent that the professor-expert's elaboration of ideas continued to develop throughout this project, and indeed new practices and theory continue to emerge within the discipline. Until the participant researcher could resolve any perceived (not necessarily real) inconsistencies in the experts' logic the task was not completed. Until the professor-expert commented upon the knowledge structures as depicted by the teacher-researcher and as action research demands, the task was not finished. It was recognised that the task would need to be limited to the sorts of statistical ideas that were suitable for introductory statistics teaching programs. The cessation of the process of unpacking expertise was necessarily arbitrary and driven by time constraints. It is not possible to know all that another knows, nor therefore to fully extract its meaning. Conclusions were expected to identify unresolved issues.

2.5 Positioning in relation to theories of coming to know the world

There are various theories of how it is we *come to know*. Several of these have, as integral parts, ideas about the structure of knowledge and these ideas in turn can be linked to the use of knowledge structures in teaching and learning. At this stage, the examination of these theories was from a positioning perspective, rather than an attempt at resolution of what the various theories had to say about coming to know. It was important to identify those aspects of theory that had become embedded in the psyche of the teacher-researcher. Those aspects formed the starting position from which the grounded theory of learning was to emerge. Hence it was important to clarify personal knowledge of them and what had been embraced within them.

2.5.1 Positioning the researcher

From the beginning, there was a gravitation to aspects of what was then discerned as two broad sets of theories, Cognitive Psychology and Personal Construct Psychology (PCP) (Kelly, 1955a; Kelly, 1955b). For the researcher, there was a sense, initially not well articulated, that they were different. For some, Kelly's PCP is viewed as a cognitive theory (Rychlak, 1983), whilst others suggest that Cognitive Science pays less attention to emotions and individual differences than does PCP (Gardner,1985). Similarly, Stevens (1994) suggests that cognitive theories emphasise the cognitive or intellectual nature of processes rather than 'the undifferentiated intellectual/emotional/conative means of man interacting with the world'. Irrespective of how each set of theories could be classified, what was absorbed from each theory was not exhaustive of the theories; at times there were similarities in ideas, and at times different aspects of each were drawn upon. Those ideas which were perceived as influential were the ones which were taken up for discussion.

2.5.2 Contributions from Kelly's Personal Construct Theory

It was Kelly's theory that was pivotal in my recognition that I could act as a practitioner of Psychology. It provided me with a way of being a Psychologist or therapist, a way of working with people. One of the central ideas that connected with some sense of personal philosophy was that of constructive alternativism. As Kelly explained:

> The theory is based upon the philosophical position of constructive alternativism, the notion that there are many workable alternative ways for one to construe one's world. The theory itself starts with the basic assumption, or postulate, that a persons' processes are psychologically channelized [sic] by the ways in which he anticipates events. That is to say human behaviour may be viewed as basically anticipatory rather than reactive, and that new avenues of behaviour open themselves to a person when he reconstrues the course of events surrounding him. (Kelly, 1955b, p. 3).

My understanding of this was to view each of us as seeing the world through our own coloured glasses, these glasses representing our construct system. One implication of this is that, 'what we observe is not nature itself, but nature exposed to our method of questioning' (Heisenberg, cited Lincoln & Guba, 1985). In this manner the richness of individual differences in how people see the world is understandable:

> What we think we know is anchored only in our assumptions, not in the bed rock of truth itself, and the world we seek to understand remains always on the horizons of our thoughts. To grasp this principle fully is to concede that everything we believe to exist appears to us the way it does because of our present construction of it. Thus the most obvious things in the world are wide open to reconstruction in the future. (Kelly, 1977, p. 6).

In his *Individual Corollary* Kelly states: 'Person's differ from each other in their construction of events' (p. 4). Comprehending how an individual (e.g. student or colleague) sees the world involves understanding the nature of that person's construing, and his/her system of constructs. Kelly (1955b) described personal constructs as follows:

> Each personal construct is based upon the simultaneous perception of likeness and difference among the objects of its context… Each construct is, therefore dichotomous or bipolar in nature…when a person finds his personal construction failing him, he suffers anxiety. When he faces an impending upheaval in his core structure he experiences threat. (p. 4).

Many people would be familiar with individuals using a construction of 'good' versus 'bad' on some of life's events. Kelly says of the construct sytem:

> A system implies a grouping of elements in which compatibilities and inconsistencies have been minimized. They do not disappear altogether, of course. The systemization helps the person to avoid making contradictory predictions…Within a construction system there may be many levels of ordinal relationships, with some constructs subsuming others and those, in turn, subsuming still others. When one construct subsumes another its ordinal relationship may be termed superordinal and the ordinal relationship of the other becomes subordinal. (Kelly, 1955a, p.40).

When a person anticipates the world, the subsequent experience of the world may cause constructs to be challenged and affirmed or denied. When the viability of the construct system is challenged, there may be a variety of emotions leading to a reconstrual of events. For a therapist, or for that matter a teacher, to work with an individual involves working with that individual's constructions of the world.

In his two volumes, Kelly differentiates between different types of constructs, attributes thereof, and ways in which one can work with individual's construals of the world, particularly in the context of therapy. However, the theory is also applicable in education, where learning is a primary issue. Much of the teacher-researcher's work in the implementation phase can be thought of as using the Kelly's (1955b) C-P-C Cycle. Kelly speaks of the creativity cycle (C-P-C) as 'a sequence of construction which involves in succession circumspection, preemption and control' (p. 392) wherein preemption 'commits one to handling a given situation at a given time in one way and only one way' (p. 382). The C-P-C cycle involves the initial loosening of constructions and terminates with tightened or validated constructions. He says that loosening constructs 'is a necessary phase of creative thinking' (Kelly, 1955b, p. 330). A loose construct 'is defined as a characteristic of those constructs leading to varying predictions, while a tight construct holds elements in their place' (p. 329). Discussing the creativity cycle, Fawcett (1983) cites Koestler's (1978) description of the process; 'creative originality… always involves unlearning and relearning, undoing and redoing. It involves the breaking up of petrified mental structures…and reassembling others in a new synthesis'.

In speaking of learning, Kelly describes 'learning through doing' as an evolving cyclic process:

> As the learner puts his ideas to work, he is required to formulate them concisely enough to act upon them. He passes through a series of C-P-C Cycles, each cycle requiring a multifaceted initial approach to his problem, then a temporary preemption to arrive at the issue which is relevant at the moment, and, finally, the choice of an alternative and controlled action. As he repeats each cycle, a whole area of his construction system begins to take shape, and his constructs begin to fall into a tighter organizational hierarchy. Unless he meets too much frustration, he begins to develop an approach configuration which is fairly regular. (Kelly, 1955b, p. 305).

In the teaching phase of this research, the teacher used activities to create conditions whereby the student would be open to and explore different ideas (or constructions) before assisting in the tightening of constructions in the review or summary phase of teaching.

Three strategies for examining constructs will be discussed in Section 2.4. These are *repertory grids* used to elicit constructs, *laddering* techniques used to examine the structure of those constructs, and the *snake* used to examine the development of constructs over time.

2.5.3 Contributions from Cognitive Psychology

Cognitive Psychology is the study of 'coming to know' (West et al., 1991). Various texts reveal Cognitive Psychology to be an area of study that is concerned with cognition and the 'internal processes' associated with coming to know. These 'internal processes' include, for example, learning, perception, comprehension, thinking, memory and attention (Anderson, 1985; Eysenck & Keane, 1990; West et al., 1991). Through these processes, we come to this position of knowing. As the early work for this thesis was revisited and reflected upon, a question arose about these internal processes: 'To what extent did these aspects of cognition, the starting position for theorising, remain the focus of the emergent theory of teaching and learning in the statistics classroom?'

My attention has consistently been drawn to one aspect of cognitive science, namely the structuring of knowledge. From my earliest work, I have gravitated to exploring theories which incorporated the idea of structure. Artifacts from my background included a thesis in psychology (Williams, 1975) based on the work of Gestalt psychologists with ideas of a 'central organising mechanism (or gestalten)', 'wholes', 'underlying configurational processes', 'structures' and 'insight'. Much of my teaching in Statistics, Psychology and other discipline areas has incorporated strategies, developed by cognitive psychologists, which provide a structural view of knowledge and facilitation of learning through working on these structures.

In the early days of this thesis, when reading West's (1991) account of the development of Cognitive Psychology, notes made included discussion of knowledge as composed of content (state schema) and processes (process schemata). These thoughts were comfortable, reinforcing recollections from many years ago, made when learning classes in educational psychology and teacher training about Piaget (1950) and his notions of schemata. Eysenck & Keane (1990) say of schema:

> The most commonly used construct to account for complex knowledge organisation is the schema. A schema is a structured cluster of concepts; usually, it involves generic knowledge and may be used to represent events, sequences of events, precepts, situations, relations, and even objects. (p. 275).

Schemata that are present during instruction provide the core around which detail, or further knowledge, will be added. How prior knowledge is organised into these schemata is important, as schemata exert a strong influence on perception during instruction. Perception involves the construction of meaning 'massaging the new with the old within the schemata available and activated' (West et al., 1991, p. 9). Cognitive theories share with Kelly's PCP an emphasis on the structure, but in this instance, the structure as evident in schemata rather than constructs. The image I carried of schemata was that of 'coathangers in the brain' upon which knowledge was hung and linked.

Ausubel's (1960) theories of learning and of the techniques and approaches to teaching deserve special mention as they have been extremely influential on the unpacking of expertise and my teaching practices. Ausubel states, 'The most important single factor influencing learning is what the learner already knows. Ascertain this and teach him accordingly' (Ausubel, 1968, p. vi). Supporting this approach, he recommended the use of advance organisers (Ausubel, 1960) for connecting new knowledge to old knowledge by providing 'anchorage in cognitive structure for new knowledge' (Novak & Tyler, 1977, p. 78). Cognitive structure also has a place in Ausubel's theory:

> He views the storage of information in the brain as highly organized, with linkages formed between various older and new elements leading to a conceptual hierarchy in which minor elements of knowledge are linked with (subsumed under) larger, more general, more inclusive concepts. Thus cognitive structure represents a framework of hierarchically organized concepts, which are the individual's representations of sensory experience. (Novak & Tyler, 1977, p. 25).

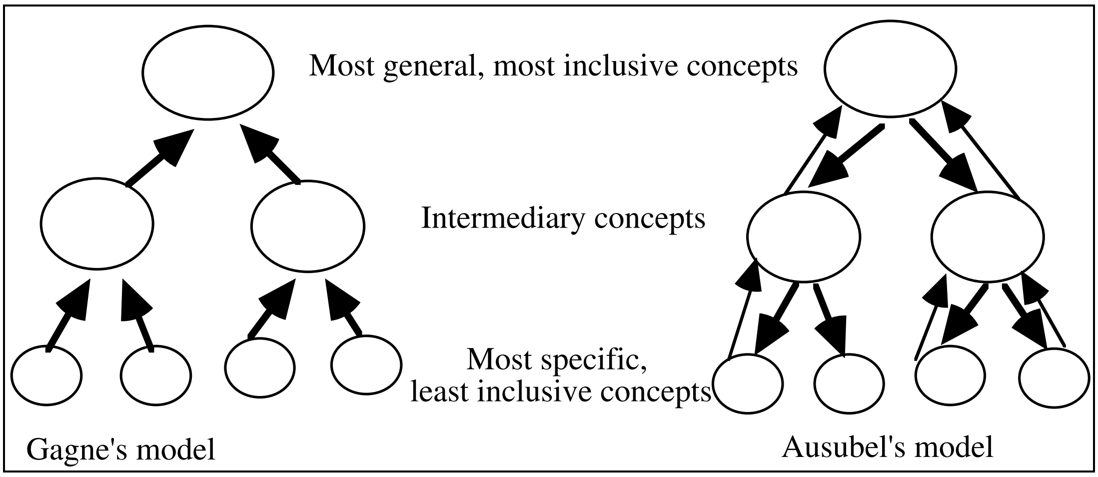
Accompanying this theoretical perspective, Ausubel's approach to teaching and learning is one which is considered to be a 'top-down' or 'general-to-advanced' approach to learning (Van Pattern, Chao, & Reigeluth, 1986). Visual/spatial strategies such as Gowin's Vee heuristic and concept mapping (Novak & Gowin, 1984) are used to support the development of meaningful learning which, in Ausubel's theory, occurs 'when new information is linked with existing concepts' (Novak & Tyler, 1977, p. 25). These will be exemplified in section 2.7.

Another theory of learning was espoused by Gagne (1977). Much of Gagne's learning theory, signal learning, stimulus-response learning, verbal association and discrimination learning is based on stimulus-response (S-R) theory (Novak & Tyler, 1977). His theorizing on rule-based learning and problem solving is similar to Ausubel in that it has a focus on the hierarchical sequencing of concepts. However, the implications drawn for the later sequencing of instruction contrasts with that of Ausubel. Gagne's theory of instruction:

> …postulates that learning is best when we move from mastery of the smallest conceptual units to the more general and more inclusive. (Novak & Tyler, 1977, p. 125).

The 'bottom-up' or 'parts-to-whole' (Van Pattern et al., 1986) approach to teaching reflects how much of Mathematics and Statistics is taught. This may be because mathematics and statistical knowledge is perceived to be built sequentially, starting with basic skills and developing into more elaborate mathematics. This approach is diagramatically contrasted with that of Ausubel in Figure 2.2.

Figure 2.2 Comparison of Gagne's and Ausubel's models of progressive concept development. (Adapted from Novak & Tyler, 1977, p. 126).



However, reference to Gagne's theory is included as it represents the antithesis of my approach to teaching. Whenever I have a holistic grasp (bigger pictures can vary) of a topic and can find ways to implement an appropriate teaching strategy, then my teaching reflects a 'top-down' approach. From this perspective one may start with a problem and then proceed to develop the theory and skills needed to solve that problem as they are encountered.

There is a connection between learning theory and the structuring of knowledge, but as the top-down, bottom-up approaches of Ausubel and Gagne reveals, the one knowledge representation may be used in quite a different ways.

2.5.4 Reconciliation of theories

In the first year, writing was not to resolve the issues which arose in comparing educational theories, nor even to determine precisely what was meant, rather it was to provide a positioning statement. The aim was simply to identify personally held learning theories in relation to knowledge. In hindsight, the influential theories examined are not exhaustive. For example, through the implementation phase, it became evident that concept of developmental stages was also an influence. However, prior to the unpacking of expertise, it did not surface as a concern.

Prior to examining the strategies for representing knowledge (Section 2.7), the term *knowledge* will itself be explored. Just as recognition of different strategies of representing knowledge facilitates their usage, so the recognition of different types of knowledge may facilitate the choice of representational or learning strategies to work with that knowledge.

2.6 Types of Knowledge: definitions

The construct *knowledge*, has been increasingly differentiated (Alexander, Schallert, & Hare, 1991). Those different categorisations of knowledge that have bearing on this study are reviewed in this section.

Alexander et al (1991) contrasted the epistemological use of the word referring 'to justified true beliefs and ... for universal, or absolute, truths' (p. 317) with how researchers of cognition viewed knowledge:

> For researchers in the field of cognition and literacy… knowledge refers to an individual's personal stock of information, skills, experiences, beliefs, and memories. This knowledge is always idiosyncratic, reflecting the vagaries of a person's own history…here, knowledge encompasses all that a person knows, or believes to be true, whether or not it is verified as true in some sort of objective or external way. (p. 317).

As has been alluded to, I set out to 'verify the truth', find the facts, the concepts, to identify an expert's structure of knowledge and I found instead, not a truth, but a construction of knowledge and in so doing identified with a constructivist view of knowledge:

> For the constructivist or naturalist, knowledge is composed of those constructions about which there is some consensus (or at least movement toward consensus) among those competent to deal with the substance of the construction. It is the most sophisticated and informed construction available but it is a construction subject to reconstruction with the

> addition of new data inputs, new claims, new concerns, new issues, or new circumstances. (Lincoln ,1992, p. 381).

It was dealing with these concepts of consensus and construction that proved most problematic. Clearly, practitioners of statistics operate from a position of absolute truth in many matters. Yet it was evident throughout the study that there are many constructions of statistical knowledge.

Implicit and Explicit Knowledge

The guiding directive for this thesis was for the teacher to translate *implicit* knowledge (not readily verbalised knowledge which may also be referred to as *tacit* knowledge) into *explicit* knowledge, so that the teacher could understand her own as well as her student's conceptions and skills. Alexander et al (1991) comment upon the fluidity of knowledge, whereby it is possible for tacit knowledge to 'resurface' from the subconscious and for explicit knowledge that has 'been the object of thought to disappear into the subconscious.' (p. 325). Bain draws on the research that examines how it is that knowledge becomes implicit:

> Extended practice in the application of knowledge appears to change the solution from one in which there is a slow deliberation about which concepts and principles may be relevant and how to use them, to fast, intuitive recognition of the relevant categories and methods (Anderson & Loynes, 1987). This not just a quantitative shift in speed of processing; it is a change in the underlying cognitive structure. A consequence of this restructuring is that some of the resulting knowledge may not be expressible - it becomes *implicit.* (Bain, 1990, p. 121).

Anderson (1987; cited Bain, 1990) described 'the processes by which *conceptual* knowledge ('what' and 'why') is gradually converted into a *procedural* form ('how'). Evidence also suggests that 'advanced learners acquire complex knowledge without being able to describe the principles upon which it is based' (Bain, p. 121). This conception of knowledge, with a focus on 'how, what and why' has been central in the explanation of why the learning of Statistics has been so problematic. In the next section, the conceptual/theoretical and methodological components of knowledge will be examined using the Gowin's Vee heuristic.

Declarative, Procedural and Conditional Knowledge

The literature provides similar but different sub-classifications of knowledge. *Declarative knowledge,* actual information, is often described as knowing 'what'. *Procedural knowledge,* knowing about processes or routines, including what order to perform subtasks, is described as knowing 'how'. *Conditional knowledge* is knowing when and where to use declarative and procedural knowledge (Alexander et al., 1991).

Well or ill-structured knowledge

Knowledge has also been classified according to whether it has a *well-* or *ill- defined* structure:

> Well-structured knowledge may be characterized in several ways: it has considerable logical order; it may have the kind of theory which allows powerful predictions, not just the kind of theory which as a label for a perspective or point of view; and it may contain families or theories or other general principles which allow inference about facts, and which allow general principles to be developed from facts or observations - which ultimately allow predictions. Structure appears to accrue either through being rule governed, as in mathematics, or some combination of being very empirical and rule governed, as in physical sciences.(West et al., 1991, p. 17).

Ill-defined structure may, in one sense, be conceived of as lacking the properties of well-defined structure. It may, but need not, also refer to knowledge that is replete with complexity, a labyrinth of connections between ideas. West et al (1991) suggests that *well-defined* knowledge is usually knowledge in reference to the physical world, whereas *ill-defined* knowledge recognizes that knowledge is a construction of the mind.

Metacognitive Knowledge

In the implementation phase, the focus of the thesis shifted to examine metacognition or *metacognitive* knowledge, 'knowledge about knowledge' or 'knowledge about one's cognition and the regulation of that cognition' (Alexander et al., 1991; Flavell, 1987; Garner & Alexander, 1989). Flavell (1987) introduced different subcategories of metacognition, which have been formally defined by Alexander et al (1991) as follows:

- *self* knowledge. 'This entails individual's perceptions or understandings of themselves as learners or thinkers' (p. 328);
- *task* knowledge, is knowledge such that the learner understands that 'different types of tasks place different demands on learning or thinking' (p. 329); and,
- *strategic* knowledge 'is a knowledge of processes that are effortful, planful, and consciously evoked to facilitate the acquisition and utilization of knowledge.' (p. 333).

Although there are other distinctions made in relation to knowledge, it is this set which are used in this thesis.

2.7 Spatial strategies for representing knowledge

A variety of visual or spatial displays has long been used for presentation of technical and academic information (Bean, Sorter, Singer, & Frazee, 1986; Davidson, 1982; Diekhoff, 1982; Gold, 1984). In this work, the emphasis has been placed upon the use of visual/spatial strategies, rather than text-based strategies. These strategies are particularly effective for presentation of the 'big picture', displaying and organizing large amounts of information, revealing hierarchical, sequential and spatial relationships, including many aspects of knowledge such as names, procedures and constraints. McFarland and Parker (1990) described the value and need for spatial strategies to represent knowledge.

> Text is a useful vehicle for presenting information, but often pictures can perform the same task more succinctly. Moreover, pictures may correspond more closely to how we actually think (Buzan, 1989). Where we are dealing with complex and voluminous data, diagrams can help us to disentangle the threads of our analysis and present results in a coherent and intelligible form...By trying to construct diagrams, we can force ourselves to clarify the main points in our analysis and how these interrelate...overburdened with detail they can become cluttered. Overburdened with complexity, they can become innaccessible. (p. 201).

From McFarland and Parker's perspective the visual and spatial strategies are useful not just for representing knowledge, but for unpacking expertise. Experts' writing, lectures and discussions provide a source of expertise. The structure in the written documents can be determined by examining content listings, transition statements, introductions, summaries, indexes and highlights of main ideas. Using spatial strategies to represent the experts' written texts required that the material be reworked in a different manner, rather than simply reproducing text. When this work was undertaken with the expert, it provided a cue for further discussion, elaboration or reconstruction of ideas; thus they became useful strategic aids in the unpacking of statistical knowledge. Statistics increasingly uses graphical tools to display statistical data. As such, there is an inherent aesthetic appeal in describing statistical expertise with graphical tools.

Three tools drawn from cognitive science dominated the representation of statistical knowledge. These were:

- frames , Type 1 (Turban, 1990; West et al., 1991).
- frames, Type 2 (Turban, 1990; West et al., 1991).
- concept maps (Novak & Gowin, 1984; West et al., 1991). There are several variations of mapping procedures. More recently they have been referred to as node-link displays (Chnielewski, Dansereau, & Moreland, 1998). These include: knowledge maps (Hall, Dansereau, & Skaggs, 1992); semantic mapping (Turban, 1990); graphic organizers, networks, (Van Pattern et al., 1986); and spider maps.

Other cognitive tools influenced the process of unpacking and representing expertise. Each of these techniques generally involves asking questions about knowledge. It is this process of questioning that provides insight into material portrayed even though the associated visual technique might not be used to represent knowledge. The tools include:

- Gowin's Vee (Novak & Gowin, 1984);
- decision trees or direct representation, production rules such as 'if-then' formats (Karake, 1990; Layne & Wells, 1990); and,
- tools drawn from use in Personal Contruct Psychology such as repertory grids (Fransella & Banister, 1977; Pope, Denicolo, & de Bardi, 1990), laddering (Fransella & Banister, 1977) and snakes (Pope et al., 1990).

The techniques that have been of major influence are discussed in more detail.

Frames Type 1

Frames use grids, arrays, matrices as a framework for representing knowledge. McFarland and Parker (1990) described frames or templates, 'these are essentially data structures which contain pointers to other frames, sets of rules, graphics, sets of procedures or questions to ask. They are generally hierarchical and may contain subframes.' Frames are useful in that they provide a coherent structure within which the detail may be organised.

Seigel and Castellan's (1988) framing of *Non Parametric Statistics* (refer Table 2.1) exemplifies the use of a frame in statistics. Their frame incorporates the use of three fundamental questions which direct the choice of test statistics; 'How many groups are involved?'… 'Are the groups related? and, 'What is the level of measurement involved in measuring the response variable?'. These ideas are communicated by labelling the rows and columns. A comparable frame could be developed to include parametric tests. An additional column, not presented in Table 2.1, extends the analysis to that of relationships as well as to tests for differences.

Table 2.1: Seigel and Castellan's Framing of Nonparametric Statistics

Measure-level	*One sample*	*Two samples Related*	*Two samples Independent*	*k samples Related*	*k samples Independent*
Nominal	Binomial chi square one sample	McNemar	Fishers exact chi square two sample	Cochran Q	Chi square k- samples
Ordinal	Kolmogorov-Smirnov One sample runs test	Sign test Wilcoxon matched pairs signed rank	Median test Mann-WhitneyU Kolmogorov-Smirnov Wald-Wolfowitz runs test Moses test	Freidman two-way ANOVA	Extension of median test Kruskal Wallis oneway ANOVA
Interval		Walsh test Randomisation test (related)	Randomisation independent 2 samples test		

An adaption from Seigel & Castellan (1988, p. inside back cover).

Whilst this is a use of framing in statistics, the professor-expert views statistics as having moved away from the dichotomy of parametric versus non-parametric statistics. However until statistics can be re-visualised, the historical structuring of statistics into parametric and non parametric statistics will remain a powerful force shaping its teaching.

Other partial frameworks within which the teaching of statistics is structured include: descriptive versus inferential statistics; and, univariate versus multivariate statistics. Textbooks (Hindle, 1993; Howell, 1995) and reference manuals for statistics software (see for example SAS[1], 1989 and JMP[2]), particularly those written for social or health scientists often provide partial frameworks to aid practitioners in their choice of statistical technique.

Frames Type 2

Frames type 2 are distinguished from Frames type 1 in that the slots for type 2 are completed through some law-like principle. Table 2.2 provides an example of a Type 2 frame constructed from an analysis of an expert's text and curriculum. The rows represent features of data which are of interest for virtually all distributions. These features include the probability function (which determines the shape of the distribution), the centre (mean), spread (variance) of the random variable and samples drawn from the population. The column 'general form' provides 'law-like principles' for determining the contents of the cells for the specific random variables, in this instance the binomial and poisson random variables.

When using this representation with introductory statistics students, the aim is to combat students' sense of 'yet another formula'. This frame was is with the intent of revealing the similarity and difference in approaches for calculating features such as the

---

[1] Registered trademark, Product of the SAS institute.

[2] Registered trademark, product of the SAS institute

mean or variance for each distribution. The formula used for the *general form* may be used for used by both the binomial and exponential distributions (refer Table 2.2). However algebraically equivalent but simpler formulae are also provided in the last two columns.

Table 2.2: Frames type 2 - Discrete probability distributions

	General form	Binomial (k, π)	Poisson(λt)
Defining characteristics, *assumptions,* *context*		• k is the number of independent trials • 2 possible outcomes for each trial • π is the probability of a successful event • π is a constant unknown parameter	• λ is the rate of occurrence • in some dimensional unit such as time (t) • λ is a constant unknown parameter
Probability function *P(X=x) or p_x*	Distribution of p_x over the values of x	$\binom{k}{x}\pi^x(1-\pi)^{k-x}$ for x=0,1,...k	$\dfrac{e^{-\lambda t}(\lambda t)^x}{x!}$
Sample Mean $\bar{x}$	$\bar{x} = \frac{1}{n}\sum_{all x} x$	$\bar{x}$ is used as an estimate for E(X)	$\bar{x}$ is used as an estimate for E(X)
Population Mean $E(X) = \mu_x$	$\mu = \sum_{all x} x p_x$	$k\pi$	λt
Variance $Var(x) = \sigma^2$	$= \sum_{all\ x} (x-\mu)^2 px$ $= E[X^2]-[E(X)]^2$	$k\pi(1-\pi)$	λt
Assessment *of fit*	• visual (eg. barchart) • χ^2_{GOF}	Same as for general case	Same as for general case

Concept maps

Concept maps are 'visual representations of meaningful relationships or linkages, termed propositions, and two or more concepts' (Novak & Gowin, 1984). These techniques have been used by various researchers to gain insight into the composition, structure and usage of experts' expertise (Novak & Gowin, 1984) and, more recently, student understanding of statistics (Roberts, 1996).

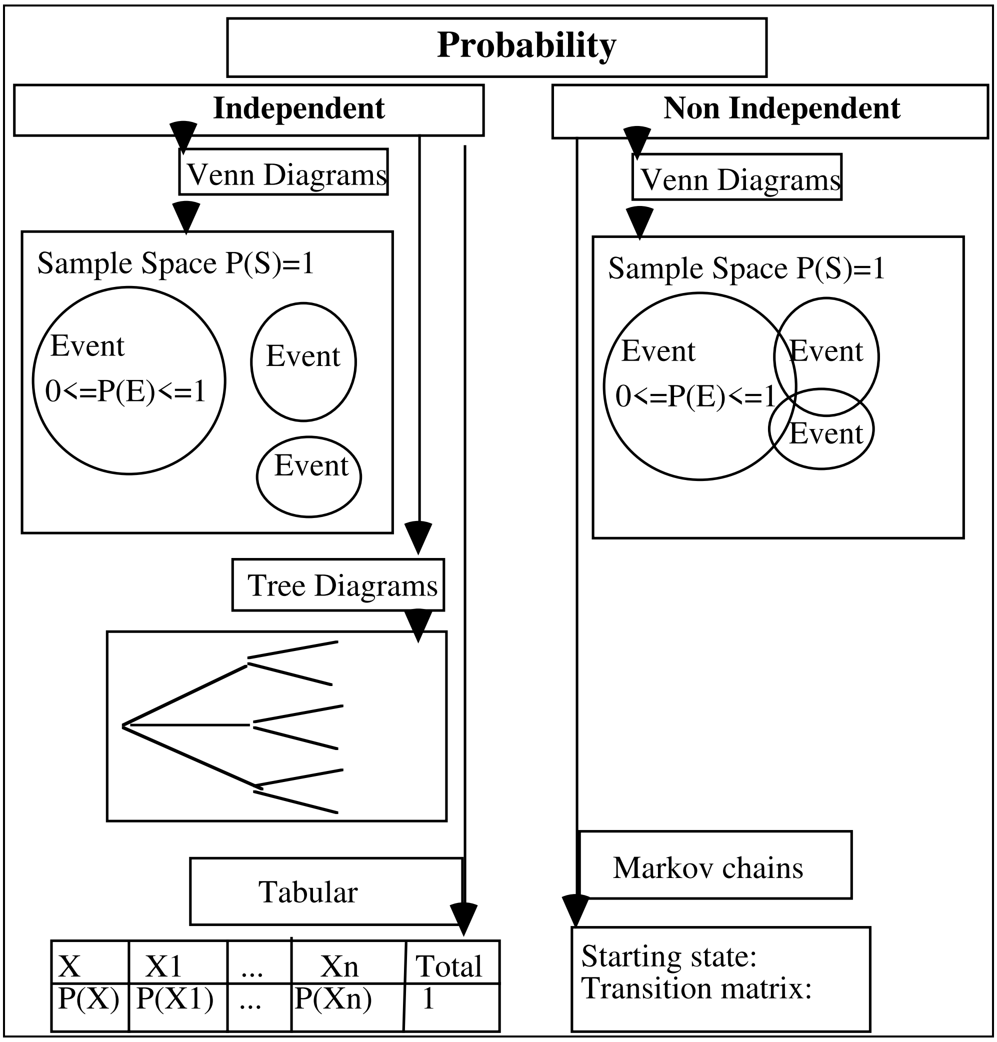
The steps involved in the process of drawing concept maps in the Novak studies included:

- identifying the concepts that are embodied in the discourse or text;
- describing the links (propositions) between concepts; and where possible,
- ordering the concepts from the most general at the top of the map to the most specific at the bottom.

What was sought in the mapping undertaken in this work were maps that were hierarchical, or alternatively, maps that retained a sense of the most important or central concept. However, when working with students, the concept maps took on a variety of forms. An example of a concept map drawn by a student in an introductory statistics course is presented in Figure 2.3. Although the student has not indicated the nature of the links between concepts this student has provided a reasonable identification of concepts and hierarchical structure. Presented in this manner it can be readily seen that the student has confused the concepts of mutually exclusive events (illustrated by the non-overlap of the two events in the Venn diagram) and independence. These ideas would need to be clarified before a more appropriate map could replace it. Just as this student did not supply a label for the link, the researcher often found it difficult, when undertaking analysis of statistical experts' texts, to determine what the nature of the link is between concepts.

Students produce various forms of maps. Different versions may or may not include a hierarchical ordering and may or may not name the links. These maps may take on a different structure to the traditional concept map. The maps may represent: a chain of concepts or *network* of events; or an hierarchical ordering of events. The mapping may be as the mind so conceives.

Figure 2.3 Student concept map of Probability concepts



Note: (Student STAT131, 1998)

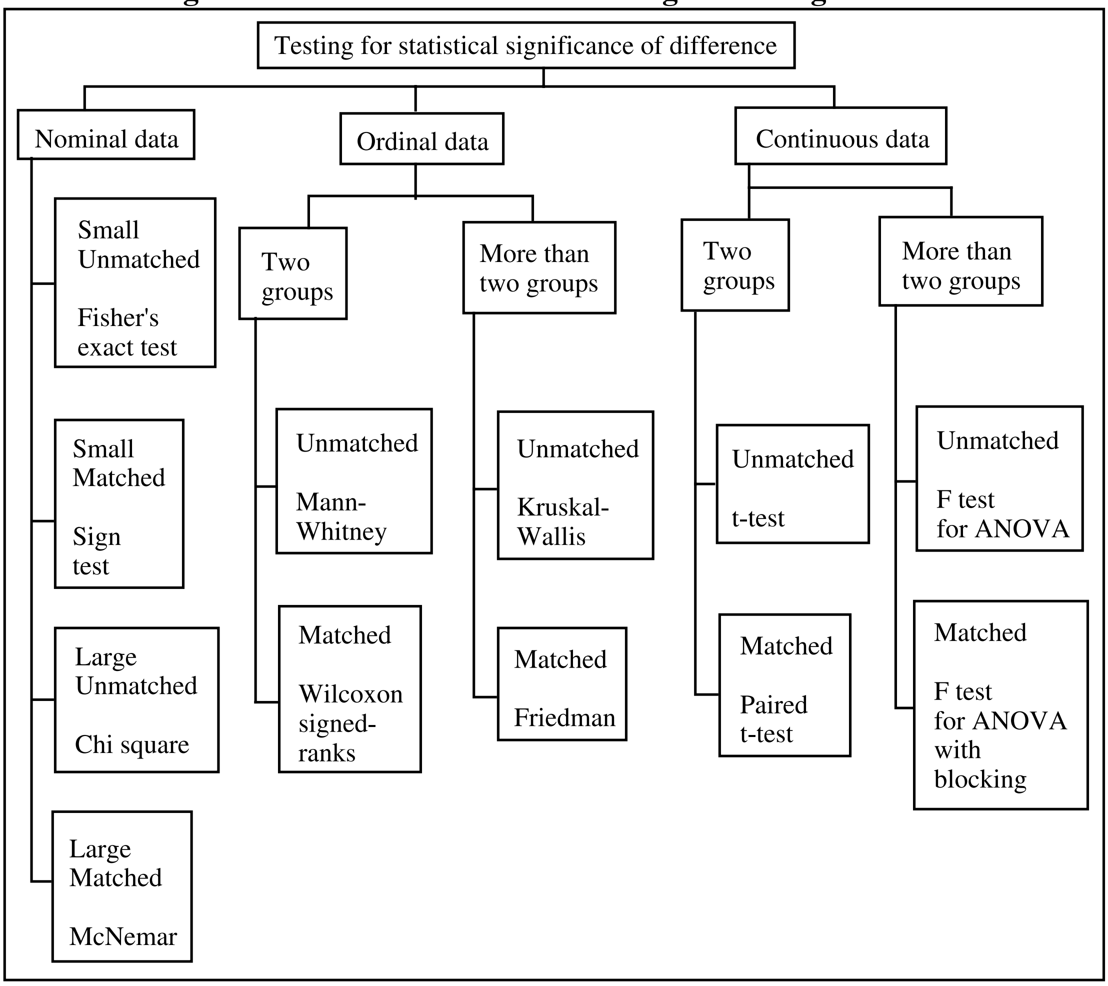
Semantic networking

When the nature of the linkage is mapped, the term semantic networking rather than concept mapping is used. Semantic networks may be produced through replacing the linkages or propositions between the concepts with terms to express the nature of the relationships. There is an extensive repertoire of ways in which links can be classified (see for example Diekhoff, 1982; Fisher, 1990). In studying student comprehension and recall, Dansereau (1978) identified different types of links between concepts. A link may reveal a hierarchical structure with type or part relationships (eg. a daisy is a *type* of flower, while a petal is *part* of a flower). Links may reveal a chain structure with procedural or causal relationships expressed (eg. a lack of water *causes* the daisy to wilt) or a link may be descriptive (eg. the daisy *is* yellow). In this work, one of the dominant organising strategies involved identifying similarities and differences, temporal steps, and similarity of function. These were often used to describe the links between experts' concepts.

Decision trees

These are trees searched in an orderly manner on simple problems. At each node or branch in the tree there is a question, the answer to which determines the next branch to be taken. A decision tree provided to graduate health students (Hindle, 1993) is presented in Figure 2.4.

Figure 2.4 A decision tree for selecting tests of significance.



Note: Adapted from Hindle (1993, p. 11).

The decision tree in Figure 2.4 directs students to decide whether the data are nominal, ordinal or continuous (ignoring discrete data); whether one, two or more groups are involved; and, whether the groups are matched or unmatched. By taking the appropriate path, students are provided with a statistical test to conduct on the data. There are, of course, difficulties with this approach, as it oversimplifies the statistical analysis. Students still need to determine whether the various analyses suggested at each branch are appropriate. That is, they still need to confirm, although it is often unstated, that the various underlying assumptions are justified.

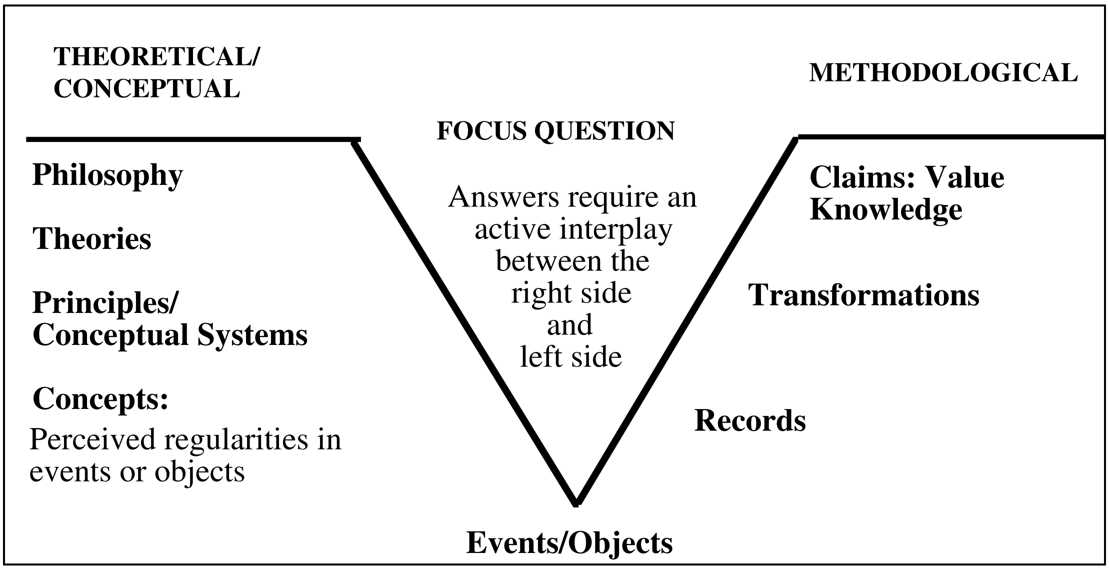
Gowin's Vee

Gowin's Vee (Novak & Gowin, 1984) distinguishes between thinking (conceptual) and doing (methodological). Knowledge is seen to begin with the observation of regularities in events surrounding some focal question. These events are placed at the bottom of the Gowin's Vee. One arm of Gowin's Vee represents 'the doing'. That is, it involves the selection of events; the recording of observations about events; and, the transformation or organisation of these records so that answers to the focal question may be generated. The final component of doing is making knowledge claims, that is, selecting the likely answer to the question. The thinking component arm of Gowin's Vee required experts or students to elucidate:

- what they were observing (the concepts for which they are to form linkages and which they are to record);
- the nature of the conceptual systems (or principles) which indicate how events appear to behave;
- how these conceptual systems and principles fit within appropriate theories which delineate why events occur as they do; and,
- the philosophical perspectives underpinning the theories and principles.

The Gowin's vee heuristic was invented to illustrate the conceptual and methodological elements that interact in the process of knowledge construction or in this case in the analysis of lectures or documents presenting knowledge, illustrated in Figure 2.4.

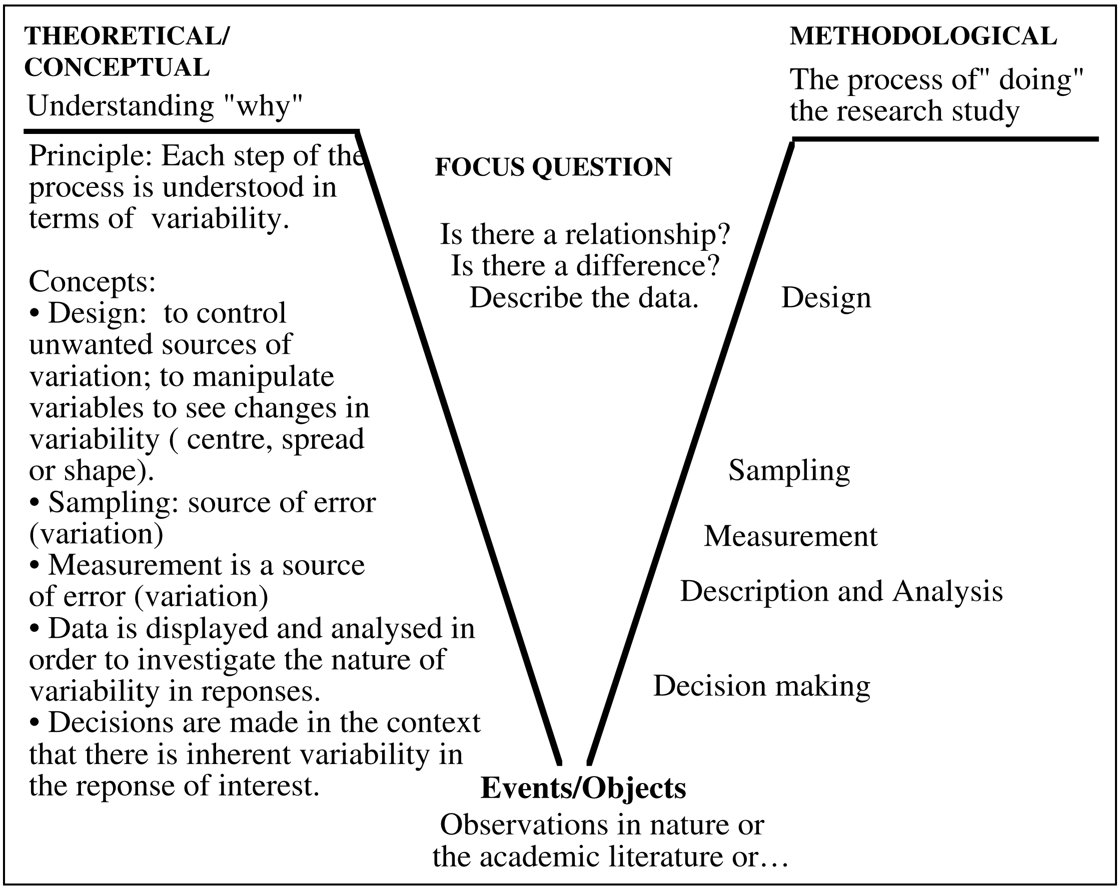
Figure 2.5 Gowin's Vee heuristic (Novak & Gowin, 1984, p.3).



Gowin's method involves the asking of five questions in order 'to unpack the knowledge in any particular field.' These five are: 'What is the telling question?', 'What are the key concepts?', 'What methods of enquiry (procedural commitments) are used?', 'What are the major knowledge claims?' and, 'What are the value claims?' (Novak & Gowin, 1984, p. 54).

These questions were used at times to guide the discussion with experts (and their texts); to clarify relationships between concepts; and as a template through which to interpret how knowledge was being structured into the conceptual and procedural. Figure 2.5 is a representation of basic statistical concepts based upon, but not strictly adhering to the questioning associated with Gowin's Vee technique. That some events are reported in the literature to occur under certain conditions may lead to the posing of a statistical question (the focal question). The focal question corresponds to the types of questions researchers ask, such as, for example, whether or not there are differences between groups, or relationships between variables or simply an attempt to describe what is happening. The *how*, the methodology or doing of statistics, may be considered to be the completing of each step in the research process.

Figure 2.6 Representing conceptual and methodological components of statistics



Repertory Grid Technique

The repertory grid technique is a technique used to elicit constructs about some aspect of our world and the relationships between those contructs. Unlike the other techniques it is not applied to the text, but requires the involvement of the person construing the world. The repertory grid constructed (for a seminar several years prior to the commencement of this study) by the researcher to illustrate the uses of the repertory grid technique is shown in Table 2.3. The focus for completing the grid was her variable performance in different academic subjects. There was no preference for any of the disciplines examined but there was a perceived discrepancy in the grades obtained and the perception that what was learned did not correspond to the grade. Some subjects therefore were proving to be sources of frustration.

There are various procedures for eliciting the grids see for example (Fransella & Banister, 1977; Pope, 1993). The procedure for eliciting the grid as represented in Table 2.3 involved selecting a set of subjects (the elements of the grid) undertaken by the researcher as a student. The subjects were selected from those that were frustrating through to those which evoked a positive feeling. They were chosen from a first year level through to the fourth year level of university study in the disciplines of Psychology, Computing Science, Mathematics and Statistics. These elements are assumed to be representative of the domain of discourse 'the learning'.

The technique for eliciting the constructs involved the continuing self questioning: 'Bearing in mind that I want to understand how it is that I best learn, how are these two subjects alike but in someway, different to this third subject?' This process of triadic comparison is repeated using randomly selected triads, until several different constructs have been elicited, and perhaps until the new constructs are exhausted. In this grid the first two bi-polar constructs are: whether or not the student had an appropriate background preparation for the subject, a 'good background' versus whether it was a 'new topic'; and, whether there was student control over the pace, of the subject, 'controlled pace' versus uncontrolled pace, 'helter-skelter'. Once all the constructs were elicited, as a way of ordering the poles, the positive poles were listed on the left and the negative poles on the right. (Other representations use the first elicited pole on the left and the second pole on the right). Each subject (the elements) was then rated according to whether it was high (5) or low (1) in terms of the positive construct.

Table 2.3 Repertory Grid of contructs about learning

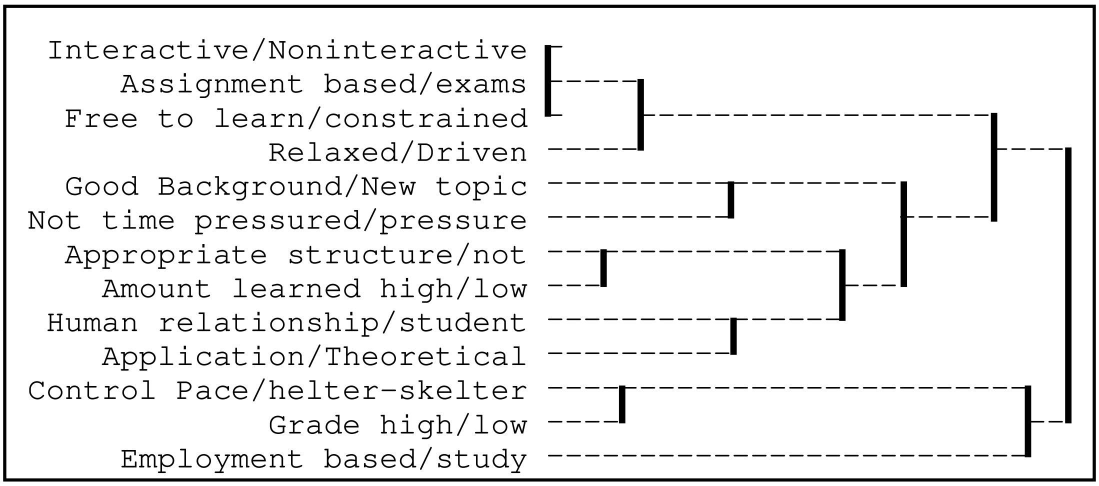
	Elements: Subjects *											
Positive Construct Pole	1 C	2 C	3 C	4 C	5 C	1 S	1 P	1 M	2 M	3 M	4 M	**Negative Construct Pole**
Good background	3	1	4	2	2	4	1	4	2	2	3	New topic
Controlled pace	1	1	4	1	1	5	5	5	3	2	4	Helter-skelter
Employment based	2	1	1	1	3	2	4	1	2	3	2	Study oriented
Relaxed	3	3	3	1	1	5	5	3	5	2	5	Ground/driven
Human relationship	2	1	3	5	1	5	5	5	3	4	4	Student
Appropriate structure	5	4	3	4	2	5	3	4	4	3	5	Inappropriate structure
Free to learn	3	2	3	1	1	5	5	5	2	3	5	Assessment constraints
Not time pressured	2	2	4	1	2	5	2	1	5	3	3	Time pressured
Application	3	1	3	4	4	5	3	1	2	4	4	Theoretical
Interactive	3	1	4	2	1	5	5	5	2	3	5	Non interactive
Assignment based	2	1	4	2	2	5	5	5	3	3	5	Exams
Grade High	2	1	4	1	1	4	3	4	4	3	4	Grade Low
Amount learned high	5	3	3	4	1	5	4	3	3	4	5	Amount learned low

* The elements were different subjects in the disciplines; computing (C), mathematics (M), psychology (P) and statistics (S).

This process of questioning "how are elements similar but different to others?" has become a characteristic mode of questioning for the participant-researcher, even though the results are not formally represented in a grid.

There are several ways to analyse repertory grids. (see for example, Bell, 1983). These techniques range from simple visual analysis of constructs and elements ordered on the mean ratings to factor analysis or cluster analytic techniques. An example of a tree diagram representing a cluster analytic solution is in Figure 2.7.

Figure 2.7 Construct Clustering



Cluster analysis groups the constructs which in terms of the patterns of ratings are most similar, although it is quite possible to produce different clusterings depending on the algorithm used and, in particular, on how the distance between clusters is measured. In this example the constructs 'interactive', 'assessment-based', 'free-to-learn' have the most similar patterns over the subjects rated. Those clusters which are linked last are least alike. In this grid the 'amount learned' clusters most closely with 'an appropriate course structure'. Interestingly, developing an appropriate structure for material is high on the teacher's agenda in this research, just as it was when she examined the courses she completed as a student, some years prior to commencing this study. The high grade construct clusters closest with courses in which the pace is under the control of the student. Unfortunately, neither the cluster analysis nor the factor analysis reveals the organisation of constructs.

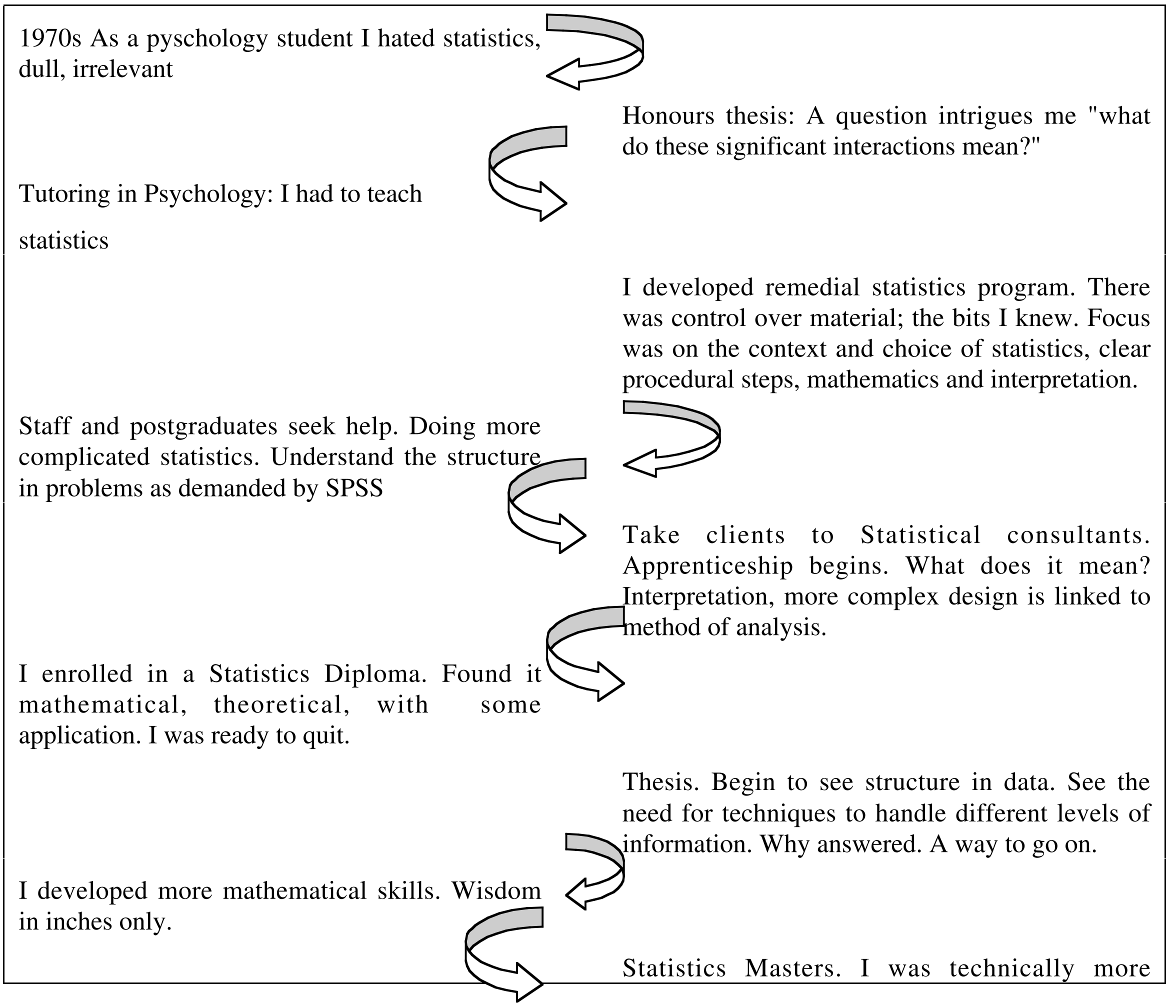
Laddering

Techniques such as laddering (Fawcett, 1983; Fransella & Banister, 1977) may be used to reveal the superordinate constructs. Once a construct emerges, the opposite of that construct is elicited. The subject is then asked which is the preferred pole, and why that is so. This elicits a further construct, the opposite of which is again elicited and the subject again asked why this is the preferred concept. The procedure continues until no further construct can be elicited from asking the question 'Why?' Laddering from different initial constructs will often lead to the same few superordinate constructs.

The snake

Another constructivist tool which allows an exploration of constructs as they develop over time is the snake (Pope, 1993; Pope et al., 1990). In general, participants are asked to visualise their lives as a winding snake or river. Each turn in the river is recorded as a personal experience or incident that influenced the direction of their life. Their *life*, as in this study, may be more focussed upon their educational life, or other aspect that is of interest. To illustrate this technique, the researcher applied it to her own statistical life, weaving between experiences as a student and those as a teacher of statistics. This journey represented in Figure 2.7 reveals the importance to the researcher of identifying structure in statistical knowledge along with a need to understand, not just what test to choose when, but to understand 'Why would one develop this sort of test in the first instance? and 'Why is the test a source of revelation about the data?'

Figure 2.8 Snake segment: turning points in the researcher's statistical journey.



proficient. I could match context to problem. My apprenticeship continued. I knew how different consultants approached problems, and how they would explain.

Slowly aware: statistics is a study of variation. I understand: the criteria to judge and develop statistics; the how and why of the mathematics. I accept uncertainty in solutions. I begin to take ownership of statistical knowledge and choice in procedures, dictated by knowledge I own and understand, not what I know is done by others.

1990s. The journey continued as techniques were re-examined in the context of making a statement about variation and a search for further principles.

In later phases of the thesis, this technique was applied primarily to understand the students' journeys through statistics. As it was not always possible to interpret whether a turn in statistics had a positive or negative influence or emotion attached to it, the metaphor was modified. The instructions were for the students to draw a landscape with the inclines and peaks of the hills to represent the high points, the positive aspects of their learning and the downward slopes and valleys to indicate the low points.

2.8 The use of spatial strategies in teaching and learning

These spatial strategies are not simply mechanisms to represent knowledge. The production and use of them may be used in the teaching and learning process. Teachers can effectively use frames at the beginning of lessons to provide a structure or at the end to review work. From a teaching and learning perspective, students can be asked to develop the frame or map, thus studying the material in the process. Through the construction of spatial representations, students can examine a topic to discover the major ideas, concepts and principles. From a cognitive theorist's perspective, the learning of the ability to develop visual representations would be important as this is part of learning how to learn. For example, the mere construction of a frame can instigate the search for additional information as parts of the matrix are identified as incomplete. In particular instances these statements may all be 'truths', but the issues in using frameworks are more complicated.

The issues in using various spatial representations focus on questions such as, 'Should students be required to develop their own frames and maps?', 'Should they be assisted in the construction of the representation?' and 'Should teachers provide frameworks for students?' One potential danger of framing could be to lock students into a framework for the discipline that is non-optimal. It may be that the frame provided by the expert will yield a better representation of the knowledge, but is this more important than the student constructing their own representation? Perhaps the process of coming to realise that one's own framework is non-optimal is an important part of learning and creating ownership of learning.

If teachers are to provide a framework, when should it be provided? Knowing how and when to use a frame is not a simple task nor does it always aid in the manner we, as teachers, anticipate. One of the teacher-researcher's experiences in an introductory statistics program involved the use of a Type 2 frame similar to that provided in Table 2.2. The first time that the teacher-researcher used this framework in the introductory phases of teaching discrete random variables, the students quite obviously 'lost the plot'. Faces revealed sheer panic when introduced to the partially completed framework, even though it was used to say, 'We will examine the same features for the general case of discrete distributions and two specific examples'. The panic was such that in the next lecture an alternative approach to the topic was utilised. However, the same framework provided as a summary was relished by students suggesting that in this instance it was more appropriate as a review strategy than in the role of an *advance organizer*.

Further questioning may take place as to the nature and efficacy of any learning that results from using frameworks. Whilst spatial representations have been found to enhance knowledge aquisition (Rewey, Dansereau, Skaggs, Hall, & Pitre, 1989), these findings are not always replicated (Hall et al., 1992). When discussing the reasons why their findings were not replicated in a second study, Hall et al posed two possible sources of difference; the nature of the map's structural design, and the nature of the subject matter. Possible differences in the structure of the map which needed to be considered included whether or not the map was hierarchical, simple or detailed, and whether it had a main idea that was easily recognisable. The nature of the material included was different in its level of abstraction, complexity and familiarity and being part of the 'students' experiential repertoire'.

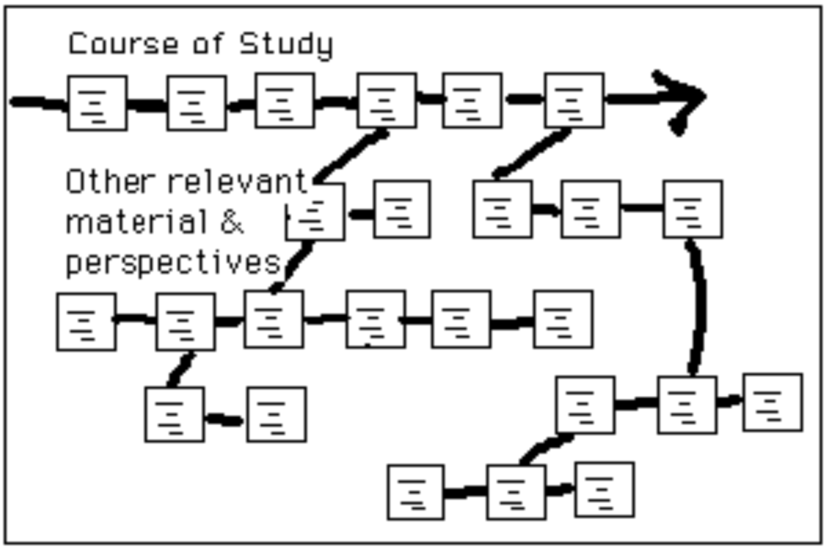
These issues relating to the use of spatial representations are addressed again during discussion of the implementation of the statistical curriculum, when more experience guided the commentary. Meanwhile the nature of knowledge as ill- or well-constructed and the implications for visual representations of expertise is examined.

2.9 Representing ill-defined knowledge

There is some contention as to whether or not it is possible to structure complex knowledge in the simple strategies inherent to the cognitive approach or as illustrated by the spatial strategies. Mathematics is often seen as a *well-defined* discipline, with a logical structure, often rule-based and procedural, closely reflecting the structure of the physical world in contrast to disciplines such as psychology which are *ill-defined*. Mathematics is based in 'verifiable fact' (or at least has an axiomatic basis) but psychology in 'movement toward consensus'. Theoretical statistics, like mathematics, is axiomatic and as such is also well-defined. However, the application of statistics often takes place where there is no theory to direct it, for example, when assumptions are violated and no statistical tool is strictly appropriate. Hence, applied statistics could be classified as ill-defined, although it may reveal more consensus than psychology. If one adopts the constructivist description of knowledge provided by Lincoln (1992, refer section 2.6) and applies it to all knowledge, then the well-fined versus ill-defined definition of knowledge loses some of its potency, as all knowledge may be seen as a construction of mind.

One approach to structuring the learning of complex material is illustrated by the *Theseus* research project (Stringer, 1993). This project used hypermedia to bring about learning. Emphasis was placed on the structuring of learning rather than instructional/teaching technique. *Theseus* allowed various kinds of processes to be explored by the student from different perspectives. Whilst the student embarked on a linear journey in terms of the knowledge to be attained, deviations in the journey were permitted to allow full exploration of the concepts from other vantage points (as illustrated in Figure 2.9).

Figure 2.9 Students journey through the labyrinth



Stringer (1993) demonstrated materials developed for the training of laboratory technicians. The approach relied heavily on student's seeing visual images of a large number of normal and abnormal human cells, and the subsequent exploration of them. Students through the medium had access to abnormal cells rarely seen under the microscope, in addition to those seen frequently. Students saw what the teacher saw as they engaged in the process preparing slides for examination, focussing the microscope, rotating the cells in space. All aspects were all readily presented in this medium. This and other processes could be explored in seconds rather than remain inaccessible or be attained only with protracted time and/or difficulty. Students were required to be active, stopping, retracing material and examining material presented in various perspectives. That is, they retained responsibility for focus of their learning. Stringer characterised the course of study as a linear path, which had various offshoots of relevant study.

Cognitive flexibility theory is appropriate theory for subject matter that is ill-defined or complex, as illustrated by the *Theseus* project. In *Cognition, Education, Multimedia: Exploring ideas in High Technology,* Spiro and Jehng (1990) discuss Cognitive

Flexibility Theory and the application of what is termed hypertexts. Hypertext can be readily produced with current technology and may be any material (text, video, auditory material etc.) which is presented in a way that allows non sequential access for exploration of concepts from multiple perspectives. Thus the learner is assisted in restructuring knowledge in many different ways and hypertext facilitates the aquisition and transfer of knowledge. Spiro and Jehng refer to the structure as nonlinear and multidimensional.

Spiro and Jehng (1990) described the advantages of what they call Cognitive Flexibility theory as follows:

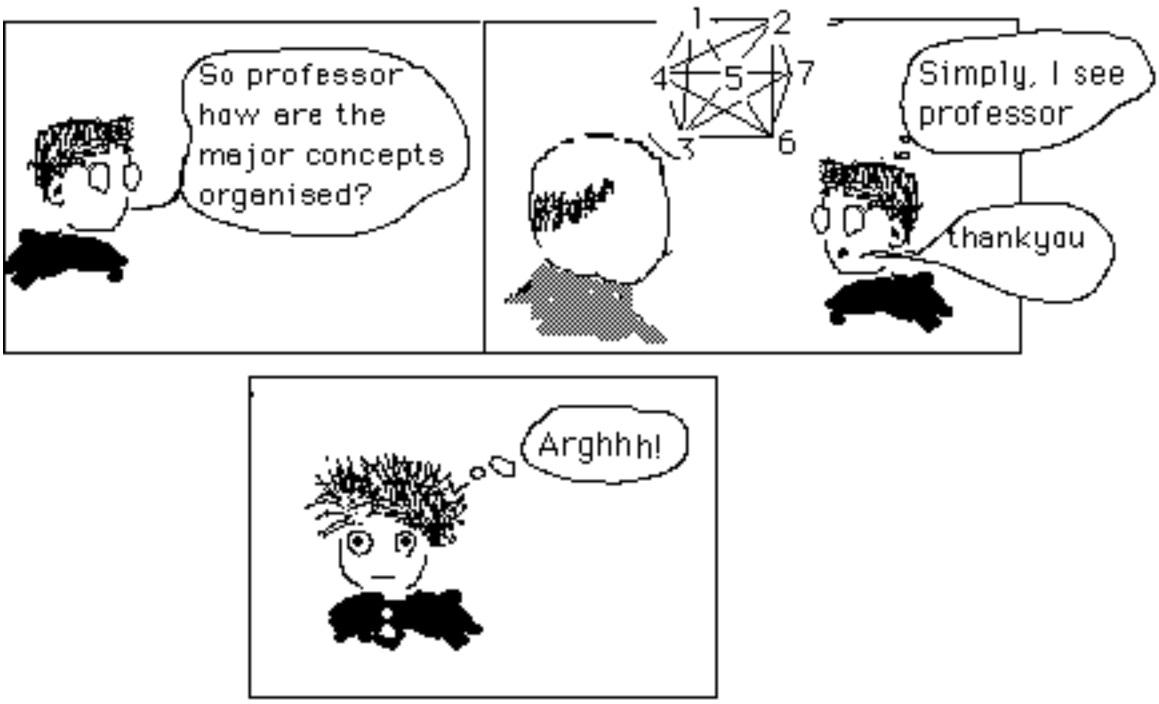
> ...by crisscrossing a conceptual landscape in many directions, knowledge that will have to be used in many ways is taught in many ways. (p. 171)

In computer-based learning packages or multimedia, non sequential access (often termed random access, although not random in a statistical sense) to material is seen to allow for the rearrangement of instructional sequences from various conceptual perspectives. Materials which are grouped to reveal similarity, may also be viewed to show dissimilarity. Material initially presented to show diversity can be viewed from the point of view of overlap. One of the difficulties in traditional teaching is that complexity is approached with incremental additions whereas hypermedium is viewed as a mechanism for viewing complexity from multiple perspectives.

Spiro and Jehng (1990) suggest that a common failing of advanced learners is that they try very hard to put knowledge into 'a plaster-cast knowledge structure.' They contend that much of the work comparing novices and experts is based on knowledge that is simple in structure or well-defined. They distinguish between the learning of simple or well-defined knowledge and the learning material which is complex. Early versions of schema/ frame/ script theories are considered to be antithetical to that which is necessary for complex learning. Early in the unpacking phase of this study, the cartoon Figure 2.10 was drawn to depict this sentiment. The expert, essentially drew a complex figure as shown to illustrate the structure of his knowledge, and shrugged 'who knows!'

In so doing the professor communicated that many individuals would not have fixed views on the structure of statistical knowledge.

Figure 2.10 Central concepts - who knows! (Diary extract, 1993)



The complexity of statistical ideas may be too difficult to capture in a simple set of hierarchical maps. As Spiro & Jehng suggested:

> Learning and instruction for mastery of complexity and application in a complex and ill-structured domain cannot be compartmentalised, linear, uniperspectival, neatly hierarchical, simply analogical, or rigidly prepackaged. (Spiro et al., 1987; Spiro et al., 1988). Yet it much too often is, and the result is the development of widespread and serious misconceptions in knowledge application (Feltovich et al., 1989; Spiro et al., 1988; Spiro et al., 1989). This is because complex and ill-structured knowledge domains are characterised by features such as nonuniformity of explanation across the range of phenomena to be covered, nonlinearity of explanation, nonadditivity following decomposition, context dependency, irregularity of overlap patterns across cases (reducing the effectiveness of prototypes and simple analogies), absence of wide scope defining features for category application, and so on. (1990, p. 168).

That there is a structure to statistics is evident. It may not be representable in simple conceptual maps or frames or in a neat hierarchical order. There is the further difficulty that 'by selecting a particular knowledge representation scheme, one predestines many of the characterisics of the knowledge that can be represented there, or even aquired from the expert' (Brule & Blout, 1989, p. 14). With the caution that Cognitive

Flexibility theory would suggest, and as is evident in chapter 3, multiple strategies were ultimately used to elicit and represent experts' statistical knowledge.

2.10 A postscript: Two philosophical approaches

Mid-thesis, an oscillation between what were now perceived as two distinct philosophical approaches (cognitivist and constructivist) played havoc. What was written one day was rewritten the next as the oscillation between the alternative perspectives ensued. The source of the oscillation was identified and described as follows:

> In the early phases of unpacking expertise there has been a search for ontological reality: 'what is statistics?' ; 'what are the fundamental concepts, the theory and logic that expert statisticians use in their construction of statistical knowledge?' Definitions of statistical concepts such as variability or measurement convey a sense of ontological reality. However, it has become apparent that these concepts can despite their definition convey different meanings and implications to different people. The search for the definitive, objective statistical knowledge structure is abandoned. It is replaced by a search for different and viable constructions of statistical knowledge that may be held by experts as evident in the literature or through unpacking exercises. (Writings, 1995).

The attitude that Still and Costall (1991) associated with Cognitive psychology could be seen to pervade initial attempts to map experts' statistical knowledge.

> Cognitive psychology has identified with the hard sciences by drawing its methods from the laboratory and its vocabulary from computer engineering ... In Science, ..., language seems to be geared to providing unambiguous and precise descriptions of reality, whether at the level of theory or observation. Once a true account is given there should be no room for disagreement, and persisting arguments can in principle be settled by looking more closely in order to arrive at a consensus. When this is done there is a perfect transparency through words to the reality they articulate. All traces of vagueness will then be eliminated. Science, on this account, aims for truth that is essentially incontestable. (Still & Costall, 1991, pp. 7-8).

The trap I had unwittingly fallen into was regarding the usage of the words, constructs and concepts/schema. On one level they were treated as synonomous, but at other levels quite differently. When the terms concepts and structure of knowledge, were used there was a search for something that existed, some truth, some common meaning of statistical language. The terms concept and constructs were used as if synonyms and yet there was an implicit awareness that they were different; hence the oscillation in writing. When concept was used it was as:

> … [a] representation of an objectively informative reality! Thus reality is seen to imprint itself upon our understanding, or at least is seen to be available for a subsequent checking against our representation. (Stevens, 1994, p. 6).

However, Kelly's use of the term construct is at variance with that of concept for Kelly's assumption is that:

> …reality does not announce itself to us as a series of directly discernible, discrete entities, but is construed as such: the substance which he construes does not produce the structure; the person does. (Stevens, 1994, p. 6).

In the book *Against Cognitivism* (Still & Costall, 1991) philosophers refer to the 'dangerous' illusion provided by selected examples that cognitivism is somewhat plausible. 'The plausibility is gained by seeming to come close to our common experience.' (Mixon, 1991, p. 29). The alternative to the cognitivists approach to understanding the development and functioning of the mind according to Bolton (1991) requires that psychology be based in 'a description of the different ways in which the world can be understood' (p. 110). This is the position to which the teacher, as researcher, had moved during the unpacking and teaching phases as focus shifted to an unpacking, not of concepts, but of constructions of statistical knowledge. The shift also meant that the unpacking was primarily of self and not other experts. Other experts provided a stimulus, but ultimately the expert was me, searching and directing the study process until coherence in the constructions of others had been found.

CHAPTER 3

Unpacking Statistical Expertise

> Knowledge acquisition begins by locating and collecting written sources of information such as manuals and references. It may use case study method to assist the domain expert in describing their problem solving and decision making… Encourage novices and experts to describe their thought processes orally and then use protocol analysis to enable better understanding of decision making…The heuristic knowledge is hardest to get at because experts - or anyone else - rarely have the self awareness to recognise what it is. So it must be mined out of their heads painstakingly, one jewel at a time…having mined these precious gems they put together knowledge bases that form the most important part of an expert system. (Feigenbaum & McCorduck, 1983, p. 94).

3.1 Introduction

Knowledge acquisition, as suggested by the expression *it must be mined out of their heads painstakingly, one jewel at a time,* is a time-consuming and difficult process. For this reason it is necessary to explore the process through which knowledge can be collected. Artificial Intelligence (AI) is the field concerned with the elicitation of knowledge from domain experts and the subsequent development of expert systems. The knowledge engineer's first task is to 'extract and make explicit human domain expert's knowledge (knowledge elicitation)' (Diaper, 1989, p. 24). Using the knowledge extracted, the knowledge engineer then develops the expert system, which is capable of making 'decisions based upon facts, rules and heuristics' identified (McFarland & Parker, 1990 p. 33). As such, general methods of knowledge engineering, as they derive from the AI field, are of use in this study.

With the tools identified to represent knowledge, as detailed in Chapter 2, the remaining task was to examine methodologies for explicating knowledge and to use those methods to unpack statistical expertise. Hence, the purpose of the research described in this chapter was:

- to examine basic principles underlying the approach adopted to elicit knowledge;
- to examine the methods of data collection and more specifically the types of questions which guided the unpacking of expertise;
- to explicate the knowledge of the professor-expert;
- to compare and contrast the professor-expert's knowledge with that of other experts; and finally,
- to unpack my own, the participant-researcher's expertise.

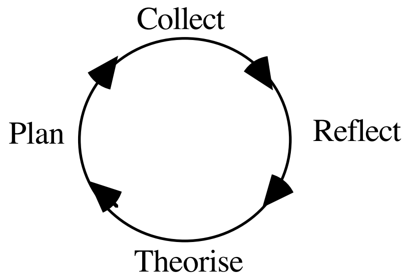
3.2 Principles of knowledge elicitation

The process to be implemented in unpacking expertise was primarily a 'naturalist' (Lincoln & Guba, 1985) and 'knowledge engineering' (Diaper, 1989) approach. The basic principles underlying this naturalist process as outlined by Diaper (pp. 54-63) and applied in the context of this study are as follows:

- *It takes place primarily in natural settings; the setting in which the knowledge is primarily used*. In this instance a collegial environment had been established and was to continue with the professor-expert and the participant researcher, co-teaching and working together on several projects.
- *It is a co-operative process* and it requires, in this instance, the co-operation between the researcher and the professor. This co-operation had already been established as part of the collegial environment.
- *It is essentially a process of research or exploration with the same class of people; permitting the asking of questions by the researcher in addition to the observation of behaviour*. Implied is that *the distinctions between researcher and subject fade away as everyone involved becomes co-researcher and co-subject*. This aspect of the process is explored more fully in section 3.2.2.
- *It is about finding tacit knowledge and transforming that knowledge into useful knowledge capable of being articulated*, although in this instance there was also an interest in the relationship between the tacit knowledge extracted and the explicit knowledge which could be readily identified through other procedures.

- *It is an iterative process following the Experiential Learning Cycle* as illustrated in Figure 3.1. The stages involve collecting data, reflecting upon it, organising the data emerging from the reflections; and planning for the next phase of the cycle.

Figure 3.1 The experiential learning cycle (Diaper, 1989, p.60)



- *It is a process that needs auditing*, or validation. The representations of knowledge elicited from the professor-expert were returned to him for commentary and modification. Findings were triangulated with knowledge explicated from a variety of experts, utilising a variety of procedures, with a third party (a second supervisor) involved in joint discussion of the materials developed between the professor-expert and the researcher. Even so, validity in this context was not to verify an objective reality, but to ensure that what was produced was in some sense an appropriate perspective of the phenomenon.

- *It is a process that produces context sensitive results* and, as such, should only be interpreted in the context of the inquiry rather than generalised to contexts other than from which the data were collected.

3.2.1 Knowledge Collection

Knowledge can be collected using one of a number of elicitation techniques (Brule & Blout, 1989; Burge, Online; Diaper, 1989). Several of these techniques were used in this study. These included:

- Interviews. Interviews were primarily unstructured, developing as the interview progressed rather than structured with a predetermined set of questions. Questions were at times direct (or closed) questions seeking specific information, such as 'What do you mean by chance?', or indirect (open) questions, which are exploratory such as 'Why would you choose to use a logistic regression instead of a discriminant analysis?' Associated with each of these questioning approaches are techniques for probing deeper (see for example, McGraw & Harbison-Briggs, 1989) and these were, at times, utilised.

- Document analysis, and the examination of other artifacts.

- Observation. Experts were often observed, without intervention, as they undertook tasks.
- Protocol analysis. The professor-expert was engaged in 'thinking aloud' as tasks were completed (see for example, Ericsson & Simon, 1984).
- Case studies or problems of special interest. These provided a context for questioning the expert.
- Critiquing. The researcher's approaches to solving problems, representing knowledge and teaching were at times reviewed by an expert statistician.

One additional technique, the joint undertaking of work with experts, permitted the use of all these other methods. In this instance the joint work with the expert involved writing papers, multimedia development, teaching, consulting and committee work. All these instances of *joint working* provided opportunities for explicating knowledge through observation, discussion and questioning.

Informal situations, such as having a cup of tea with the professor-expert or visiting scholars were often the source of ideas requiring subsequent follow-up. As Rauch-Hinden (1988) said:

> … because so much of the expert's knowledge is intuitive and subconscious, at times there is as much to be learned from ramblings as from direct questions. (p. 107).

Even when techniques such as concept mapping, framing, building semantic networks, and Gowin's Vee were not formally used to represent knowledge, the questions posed in using these strategies made it possible to address the general questions:

- What is the nature of experts' knowledge?, and
- How do they structure and use their knowledge?

This was accomplished by interrogating the various formal and informal data sources using the following subsidiary questions:

- What are the central statistical concepts constituting basic statistical knowledge?
- How and why are these concepts linked together?

- What is the nature of the links? For example, did they indicate similarity or dissimilarity, rationale or description or contrasts of some other kind?
- Are there hierarchies of concepts?
- How can the knowledge structure be represented? and,
- What is the impact of the concepts on statistical theory and statistical practice?

In an educational context, another general question was also important: 'Did the statistical knowledge as presented to students provide for coherent, comprehensive and meaningful learning?' and hence was it, as Novak and Gowin (1984) termed it, 'a good piece of knowledge.' That is, did the knowledge embody both the thinking and doing arms of the Gowin's Vee as was deemed necessary for meaningful learning or did the students learn a process without the associated understanding of the concepts involved.

3.2.2 A process of exploration with the same class of people

In adopting a phenomenological stance, that there are alternative perspectives, the intent had been to describe rather than to critique or elevate one perspective in preference to another. The description of expertise was to provide alternative perspectives as to the nature of core statistical ideas. It was considered probable that there would be other structures of statistical knowledge. Conceivably, these would vary between experts teaching within the statistics discipline and in other domains- e.g., the psychologist teaching statistics. A less detailed study of a small number of other experts was expected to supplement this exposition of knowledge. This study, however, was not concerned primarily with the commonality of structures but rather with establishing a single coherent structure that could serve as a reference for the description and identification of differences in knowledge structures. The detailed case study of the professor-expert was to provide a cohesive framework for understanding the logic underlying statistical knowledge.

In retrospect, for much of the study, the professor-expert's position was unwittingly and unconsciously accepted as providing the 'true' structure and content of statistics rather than being on comparable footing to other experts. This reflected the student-teacher, novice-expert, relationship of the researcher to the professor-expert. It further reflected the reasons for the choice of the professor-expert. He was chosen because of his ability to illuminate the logic behind his use of statistics. Recognising this, the professor's knowledge-structure was ascribed the role of reference-structure. Ultimately, the participant-researcher adopted her own perspective, albeit growing and changing, as to essential or core statistical ideas. These ideas are incorporated into the Statistical Literacy program that is detailed in Chapter 5.

3.3 The professor-expert: unpacking the knowledge

This section describes the process of unpacking and representing the professor-expert's statistical knowledge. The unpacking of the professor-expert's knowledge was a multifaceted process. The process included formal episodes:

- analysis of the professor-expert's text and course materials;
- meetings to discuss and modify maps, frames or other interpretations developed by the researcher when analysing the professor-expert's text; and,
- the writing of joint publications.

The less formal part of the process included unscheduled observation of and dialogues with, the professor-expert. This observation was at times:

- historical, with, for example, the researcher reflecting upon former roles, as a student in the professor-expert's classes, or as a student completing previous theses;
- current, the researcher observing the professor-expert's teaching, consulting activities and interacting as a colleague on joint statistical projects, joint-authoring and team-teaching.

Dialogue included any discussion about curriculum construction and any other activity that placed the professor-expert in a position of using, representing and/or articulating his knowledge.

It is presumptuous to conclude that any structures derived from any one of these sources represent the professor-experts' structure of statistical knowledge. For example, there are potential difficulties using written materials as they may reflect, not just the nature of expertise, but other constraints placed upon production of the document. Constraints may, for example, include limits in the allowable length of a document or the knowledge base and expected degree of interaction with the intended readership. These constraints may force some material to be omitted or to be included in a particular sequence and, as a consequence, the document might not adequately represent the structure of the expert's knowledge. Similarly, the inclusion of ideas in a course may be mediated by what is perceived as easy or difficult and hence suitable, say, for a beginning student. The nature of knowledge that is on the page or presented in a lecture, cannot be deemed to be the organisation of what is in experts' heads, or the method in the experts' practices. However, using multiple strategies to elicit knowledge and returning the representations to the professor-expert for commentary provided an opportunity either to corroborate interpretations or to further clarify and develop ideas.

Despite these potential problems with textual material, the structure of the professor-expert's draft text (version 1993 & 1994)[1] was used to provide a structure for presenting and discussing the knowledge elicited through this multifaceted unpacking process. The title, the contents and the chapters of the text provide insight into the expert's key concepts, structuring of ideas and how material was chunked (a term to describe how materials are collected together). By subjecting the text to concept mapping, framing (types 1 and 2), semantic mapping and by asking questions associated with Gowin's Vee techniques it was possible to identify the organising processes of the professor-expert as they were applied to the elementary material in the text.

Local concept maps were initially constructed to map ideas and terms in paragraphs. Later, this was extended to sections of the text. The aim of concept mapping was to develop hierarchical concept maps, with the most general or inclusive concept at the highest level and least inclusive at the bottom. An attempt was made to maintain the

---

[1] Reference is generally made to the 1994 version for easier page referencing. These draft texts were co-authored with one other. By the time the text was published a third author was involved.

temporal order in which concepts were encountered. However, maintaining the 'most inclusive' to 'least inclusive' mapping of concepts often required that the temporal order of concepts be reversed. Concepts within paragraphs or sections of text were often in the order, least inclusive (for example, mean and standard deviation) to most inclusive (for example numerical summaries). Hence the synthesis of the material often resulted in a departure from the temporal order of ideas. When a global map was not initially evident, the various small or local maps of constructs were then pieced together to explore whether or not a more global map could be readily constructed by the learner. Often, when this was the case, frames or partial frames or non-hierarchical maps or sequences of maps were used to organise the ideas.

3.3.1 The professor-expert's text

The title of the draft text provides the first clues as to what the professor-expert considered essential statistical ideas, *Modelling Variation and Uncertainty* (Griffiths & Stirling, 1994). With the exception of the first chapter, which provides an overview of the statistics discipline, the title and organisation of the chapters (Refer Table 3.1) suggested that the initial section of the course was extensively involved with exploratory data analysis and presentation. This chapter was then followed by ones on Data Collection, Probability and Statistical Inference.

Table 3.1 The professor-expert's sequencing of text material

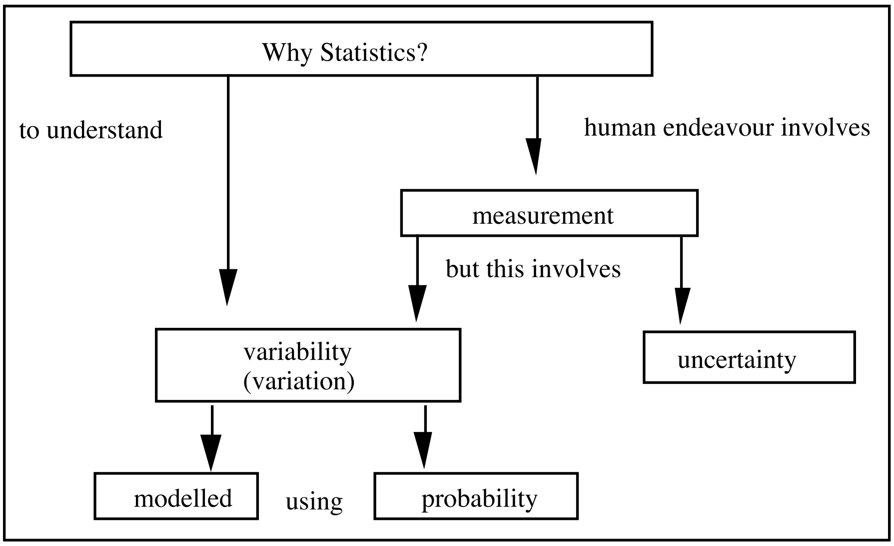
Chapter	Contents
0	Why Statistics?
1	Exploring Univariate Data
2	Exploring - Time Series
3	Exploring Relationships - Bivariate Data
4	Exploring Categorical Data
5	Principles of Data Presentation
6	Collecting Data
7	Probability
8	Random Variables - Modelling Numerical Data
9	Sampling Distributions
10	Inference for Univariate Data
11	Models and Inference for Other Data Structures

The description of the professor-expert's ideas follows the sequence provided by the text: an overview, exploratory data analysis; collecting data; probability and statistical inference.

Chapter 0 draft text: An Overview: Why Statistics?

In the initial chapter the question 'Why Statistics?' is posed (refer figure 3.2). The answer to this is that advances in 'most fields of endeavour involve measurement...But measurement involves variability and uncertainty' (Chapter 0, p. 1). What is Statistics? 'Statistics is concerned with understanding variation' (Griffiths & Stirling, 1994) that 'Statistical theory uses the tools of Mathematics and Probability' (Chapter 0, p. 1). In the second half of their course, students are reminded that we live 'in an uncertain world, we do not know in advance if it will rain tomorrow...' (Chapter 7, p.1) and that it is Probability which provides a means of quantifying and describing this unpredictability in outcomes. Probability is used to 'model the kinds of data that would be observed under different assumptions about what the world is like' (Chapter 7, p.2).

Figure 3.2. Mapping 'Why Statistics?

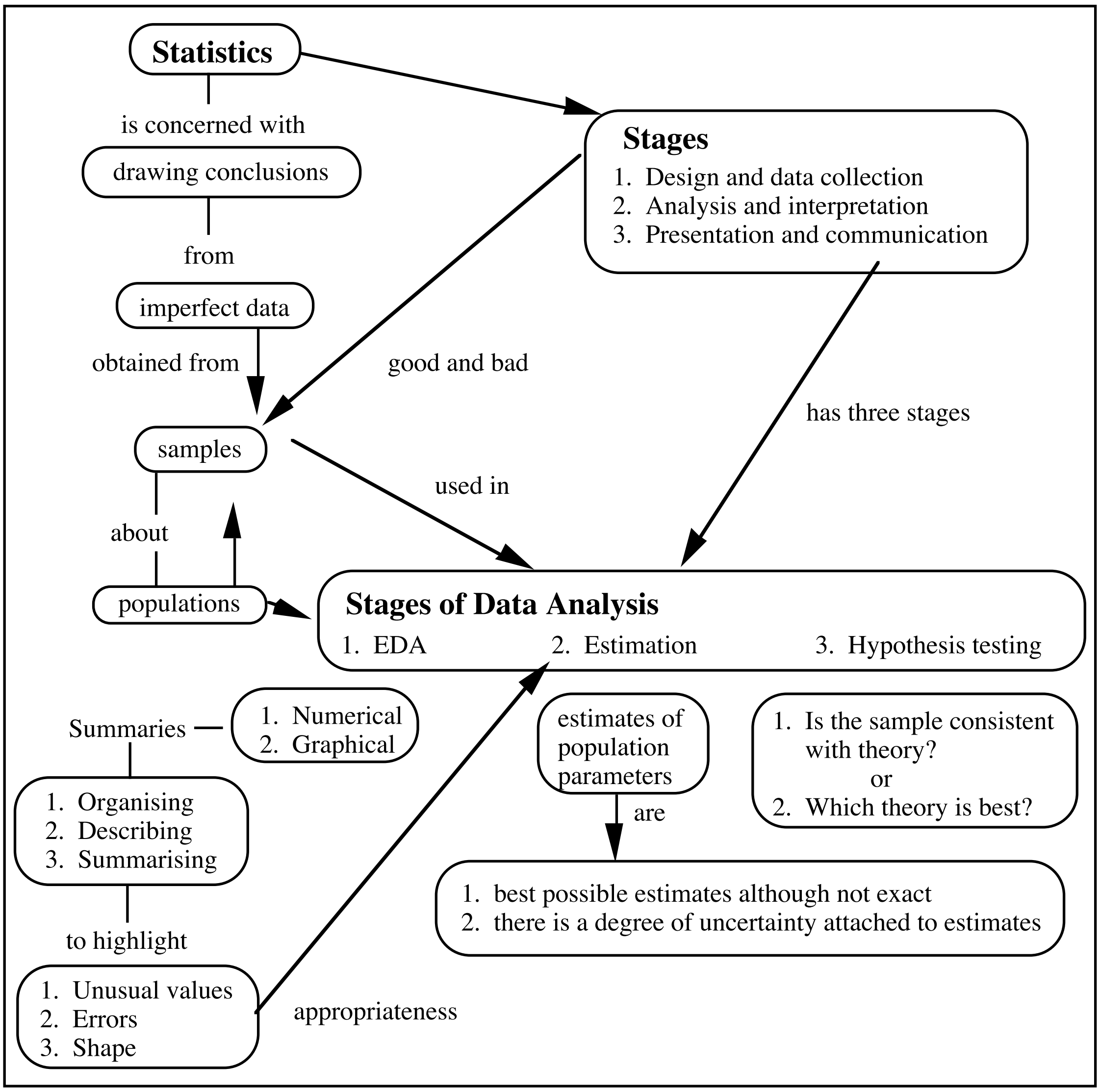


In the initial chapter, the professor-expert talks about exploratory data analysis and its relationship to the statistical endeavour and it is this theme of exploring data that is developed in subsequent chapters:

> Exploratory data analysis [EDA] should always be the first step in analysis...Exploratory data analysis reveals data structure and assists the statistician in choosing methods of statistical inference. Exploratory analysis uses numerical summaries and graphical representation to give a quick overview and to highlight features of the data of which the statistician should take particular note. Features such as unusual values and errors, or simply the 'shape' of the data may indicate that the usual and intended mode of analysis is inappropriate. (Chapter 0, pp. 1-2).

The introductory chapter introduces students to additional statistical concepts including the stages of a statistical project: design and data collection; analysis and interpretation; and, presentation and communication (Refer, Figure 3.3). Data analysis is described as two stages, exploratory data analysis (organising, describing, summarising in order to highlight features of the data) and, estimation (providing an educated guess of an unknown quantity) and hypothesis testing (seeing if a model is consistent with a particular theory). The 'why of Statistics' is again addressed with the statements: 'Statistics is the art [science] of drawing conclusions' about the nature of the world 'from imperfect data.' (Chapter 0, p. 1). The professor-expert further suggests that whether or not statistics are good or bad depends upon the design of the project and the sample drawn.

Figure 3.3 Overview map of Statistics



Complicated maps such as in Figure 3.3 provoke questions such as:

- Can the material be better structured so as to facilitate learning?
- What meanings do students derive from materials such as these?
- Are the concepts the most essential ones for the expert? and,
- How do the connections drawn from the professor-expert's written materials relate to the professor-expert's knowledge structure.

The maps represent the sense that the participant-researcher has made of the material, not the sense that the professor-expert has made. Sometimes, where the professor-expert has not provided a rationale, it has been difficult to assign links other than the temporal sequencing of text. The link 'appropriateness' is an instance that provides a rationale, a reason, addressing, at some level, the *why of statistics*. Unusual values, errors and the shape of distributions suggest something about the appropriateness or otherwise of the estimation techniques.

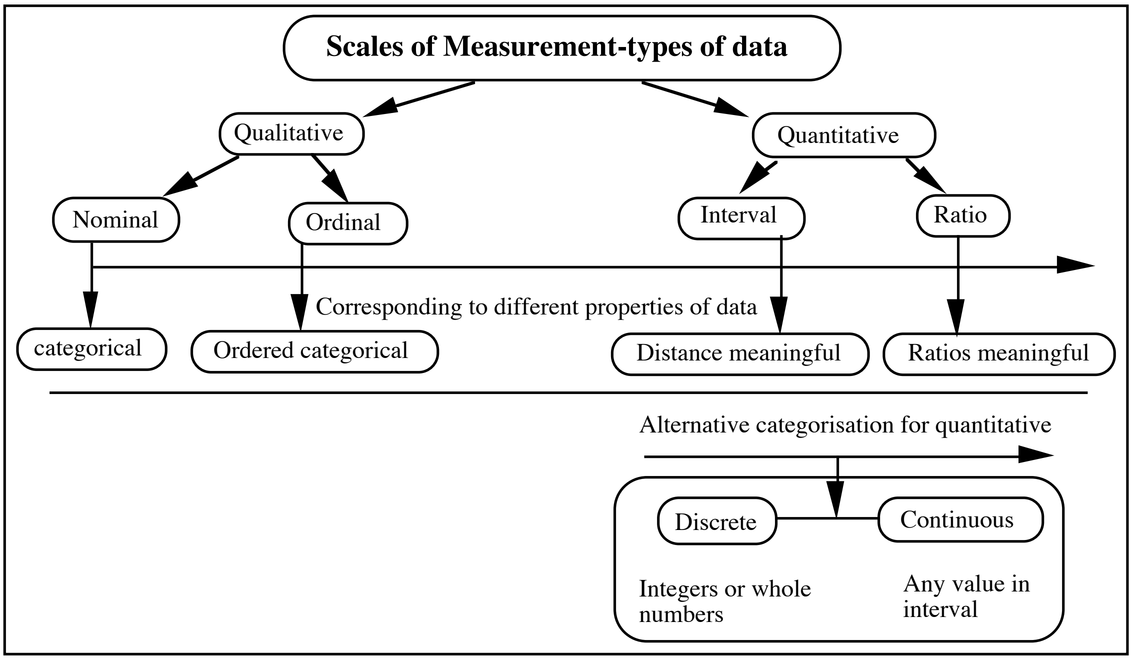
The concept of measurement was developed as in Figure 3.4. This is a fairly typical characterisation of types of measurement. It is not until one examines the means of discriminating between the various terms that differences between statisticians become evident. The preference of Hawkins, Jollife and Glickman (1992) is for the classification nominal, ordinal, discrete and continuous as they consider ratio scales to be extremely rare. The distinction drawn between these terms is problematic, as Hawkins illustrates:

> A distinction often drawn by teachers and textbooks is that 'discrete means *whole* numbers', whereas 'continuous can mean fractions and decimals'. Clearly, however, this distinction is worthless. Take, for example, the case of the 'numbers of questions answered correctly in a test of spelling'. This will be represented by 'whole' numbers, and therefore may serve to confirm the distinction being made. However, represent the same data as the '*proportion* of correct answers' and the misconception is exposed, for it is clear that 21 correct answers out of 30 is a discrete value that may nevertheless be represented as 0.7 when it is the *proportion* of correct answers that interests us. A better distinction is that discrete applies when all possible values are separated from each other by impossible values, while *continuous* applies when the possible values vary so that with fine enough measurement there would be no values within the range that were not possible. In reality, however, continuous variables are inevitably expressed by discrete versions of the continuous phenomenon. (pp. 43-44).

In the draft text the description is as follows:

> Quantitative data may be **discrete** (typically taking only integer or whole number values) or **continuous** (taking any value in some interval). Discrete variables are typically counts (for example, number of accidents, number of nonconforming items in a sample). Continuous variables are usually associated with some measurement process. (Chapter 1, p. 3).

Figure 3.4 Map of Measurement Concepts



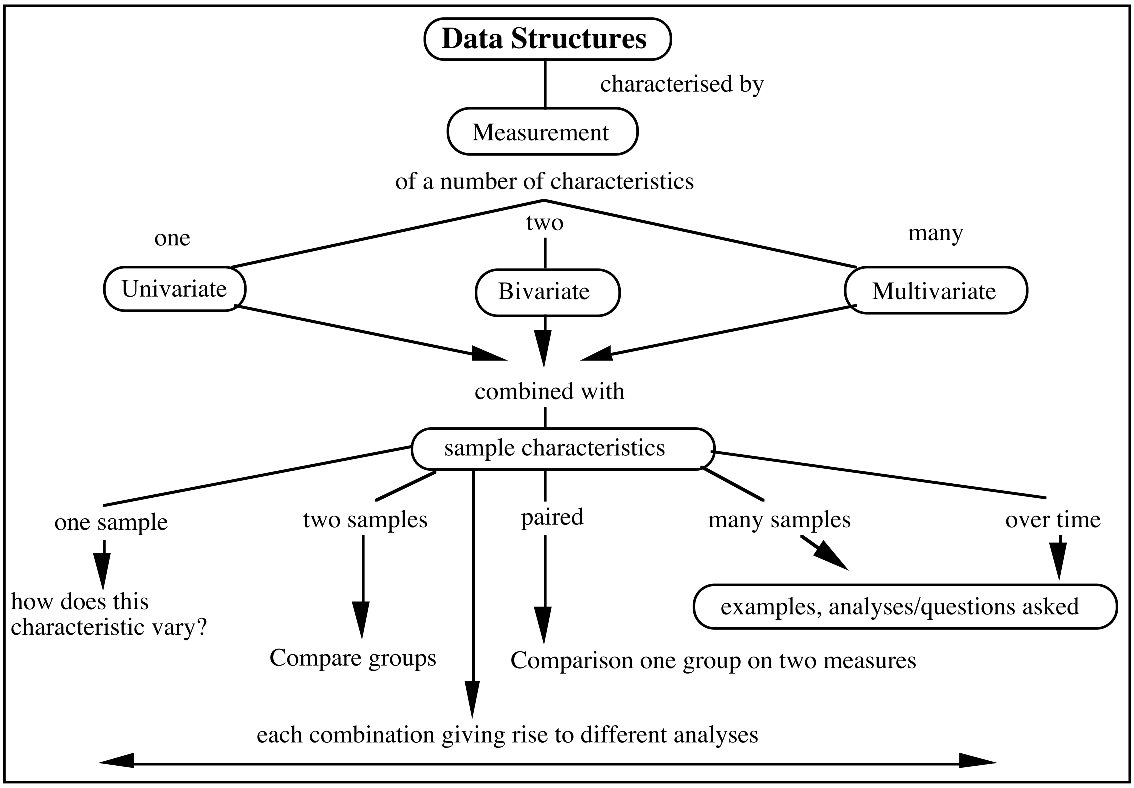
The classifications of interval and ratio and discrete versus continuous may also be examined from the perspective of cross categorisation. Interval data may be either discrete or continuous as may ratio data. Concepts associated with data structures, as in Figure 3.5, were then introduced. When the types of measurement, univariate, bivariate and multivariate data structures are combined with the sample characteristics and the nature of the question asked, a picture (refer Figure 3.5) begins to emerge of the vast range of statistical techniques that might be required to analyse data.

It was not, however, these definitions of measurement concepts or data structures that were revealing to the participant-researcher. Discussions with the professor-expert during the course of another project, *Models for Ordered Categorical Data* (Porter, 1987) had already provided the beginnings of the study of the professor-expert. The professor-expert saw more in data than did the participant-researcher. These thoughts were captured at that time:

> There are some situations in which ordered categorical analysis superficially appears appropriate for a given data set but the analysis is either wrong, best undertaken by some alternative means, or can be considered only as exploratory. These situations are as follows: (1) It is possible to perform the analysis on responses which are 'grouped

quantitative data'. In this case, quantitative responses are grouped into categories such as '1-2 days', '3-5 days' and so on... It would in these cases, be preferable to perform an analysis which utilises all the information, that is, used the raw data available prior to grouping of the 'responses', although as often happens, data may be collected in this categorical form. (Porter, 1989, pp. 12-13).

Figure 3.5 Map of Data Structures



From this interaction, another conception of data type not expounded by the definition of grouped quantitative data was perceived. Various measurement types reflected *more or less* information in the data. This simple re-construction of the definitions resulted in a change in the participant-researcher's perception of and orientation toward consulting problems. Students often seek to change a continuous measurement scale into a categorical scale, and, by so doing, reduce the level of information in the data, or alternatively they attempt to use an analysis technique that requires more information (or structure) than is inherent to the data. When students approached, with analyses planned, it was much easier to suggest based on the perception of information in data that there may be a better analysis technique, one that more appropriately used the information in the data. Also revealed in this example was the importance placed upon the response variable that the participant-researcher adopted as a defining feature of data structure as elaborated in Section 3.5.3.

Exploratory Data Analysis (Chapter 1-4 Draft text)

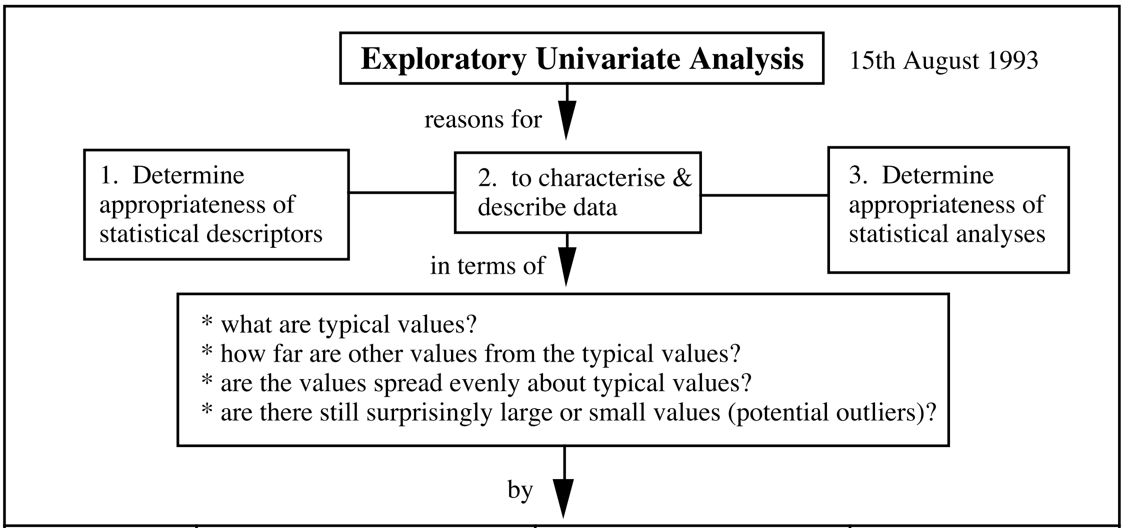
The next four chapters of the professor-expert's draft text elaborate exploratory data analysis techniques for several simple data structures. Techniques for one, two and more batches of continuous and discrete univariate data are examined as well as tools for exploring univariate batches of quantitative data collected over time (a special case of bivariate data), general bivariate quantitative data and tools for one, two and more, batches of categorical data.

Chapter 1 draft text: Exploratory Univariate Analysis: textual

An initial concept map was produced for Chapter 1 by the participant-researcher. This was then manipulated jointly by the Professor expert, one other expert (#1) and the participant-researcher in order to provide a simpler and more coherent statement of the ideas (Figure 3.6). In this one can see evidence as to 'why' the various techniques are used; to organise, to see shape, detail, location, centre, and spread in data.

The Chapter also provided the detail necessary to construct the various graphs and numerical summaries. In order to allow the comparison of two batches of data, these techniques were extended to include back-to-back stem and leaf plots, super-imposed histograms, back-to-back histograms and side-by-side jittered dot plots and box plots. The chapter finished with a warning: Statistics is not just about analysing data; appropriate methods of data collection are essential.

Figure 3.6 Exploring Univariate Data



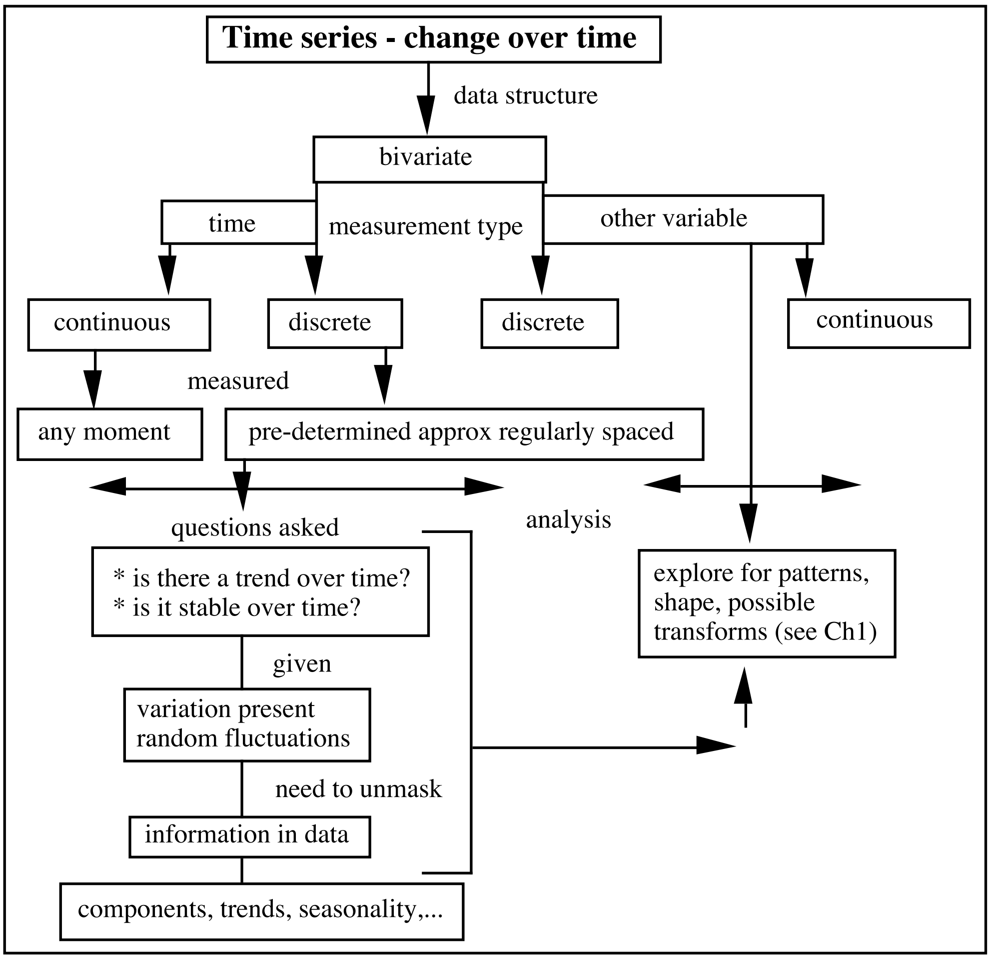
1. Organising raw data	2. Visual display (order, shape, detail)	3. Numerical summaries. Reducing data	4. Visual displays of numerical summaries
• tabular form (high and low numbers)	• dot plot (clusters & gaps range of values) • jittered dot plot (clusters range of values, density - how many) • stem and leaf (order, shape of the distribution of numbers, retains exact numbers - not good with large data sets) and for reducing data the bar chart (& histogam form) examining order, distribution binned numers (not exact), modality fine with large data sets	Location • centre (mean, median trimmed mean, mode) • percentiles, quartiles (1st, median, 3rd) & minimum & maximum Spread • about some points e.g. means(average deviation, standard deviation, variance) * between two locations (range and interquartile range) Shape • skewness • outliers	Box plots can show - Location • centre (median mean) • upper & lower quartiles, minimum & maximum Shape • symmetry • skewness • outliers
		Other • Coefficient of variation (scale free)	

But what if our data were measured on an alternative scale of measurement?
Would the characterisation of data remain the same? Transformations

Chapter 2 draft text: Exploring Change - Time Series

The second chapter examines techniques for exploring a second simple data structure, a time series. Time series was treated as a special case of a bivariate structure where measurements on one variable are collected over time (refer figure, 3.7).

Figure 3.7 Map of Time Series Concepts

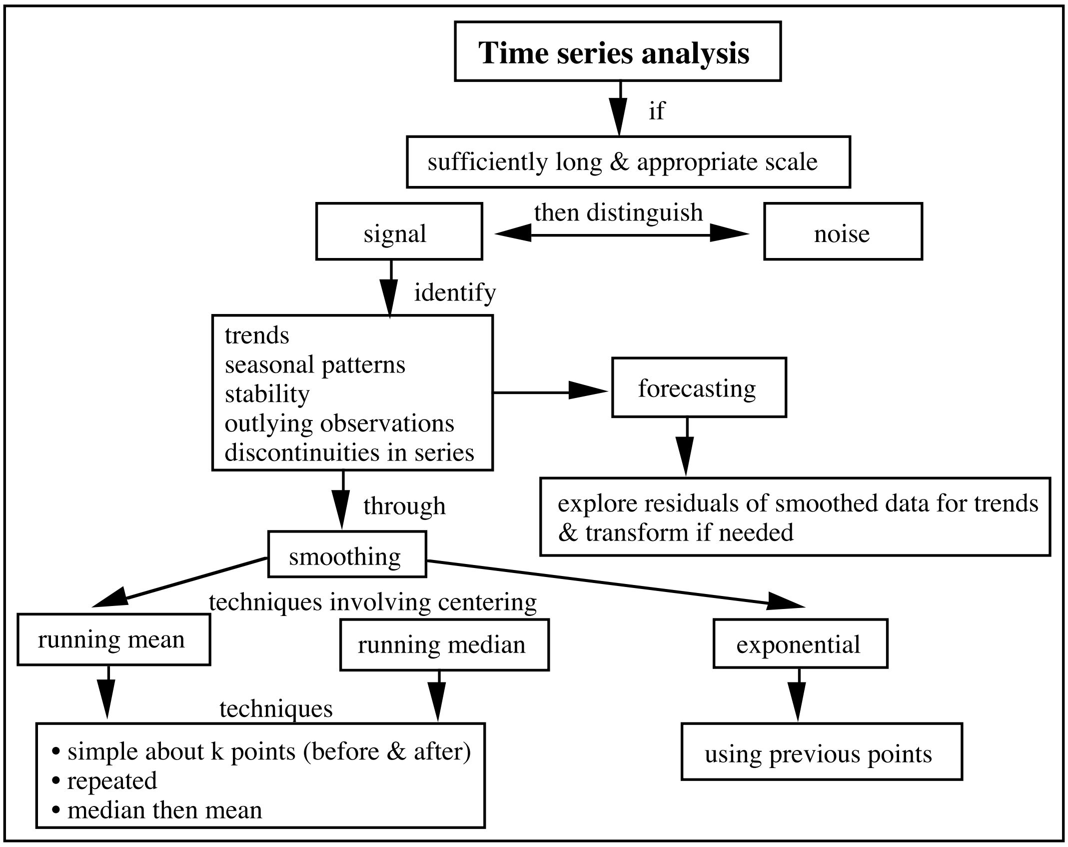


The time series techniques measure the change in a series of measurements made on one variable at different times. The examples chosen are elementary, over pre-determined, regularly spaced intervals.

The professor-expert's work again began with the application of univariate exploratory techniques to describe the measured variable. Such univariate analysis complements and informs the bivariate analyses to follow. Data were examined to see if they revealed outliers, patterns or indications that the variable might need to be transformed.

Without presenting formal details, the concepts of decomposition (refer Figure 3.8) of a series are introduced and illustrated graphically. Like so many topics at an introductory level of teaching the topic is not fully covered as the details were 'beyond the scope of this book' (Griffiths & Stirling, 1994, p. 12).

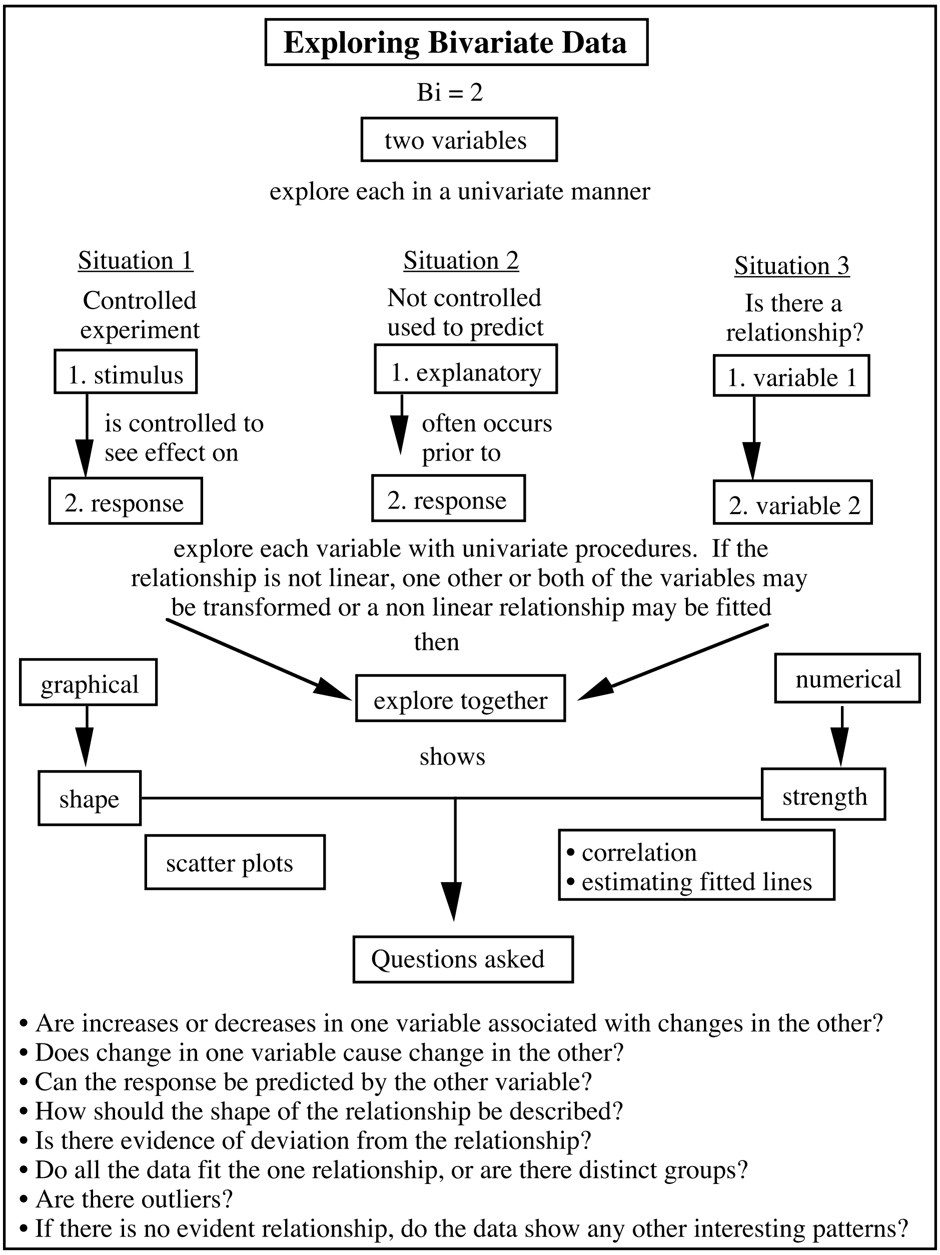
Figure 3.8 Exploring Time Series



Chapter 3 draft text: Exploring relationships

Time series serve as an introduction to the more general case of bivariate relationships encountered in the third chapter. Exploratory techniques for this type of data are examined having firstly drawn the distinction between causality and the associated controlled experiment, and association (or relationship) as mapped in Figure 3.9.

Figure 3.9 Exploring relationships



The chapter includes other subordinate concepts relating the properties of *Pearson's r* to the scattergrams of bivariate relationships, indicating how a correlation coefficient of a certain size and sign (positive or negative) may arise from data. It also demonstrates the impact of influential points and outliers on *Pearson's r*.

Chapter 4 draft text: Exploring Categorical Data

In the earlier chapters the term *data structure* has been used to describe the data analysed. Thus far, univariate and bivariate analyses have been examined with one and two batches of data. The expression *data structure* was not used to describe the data in Chapter 4 (Griffiths & Stirling, 1994) although it appeared to be implied. Lacking the explicit reference to *data structure* the questions arose:

- how does the professor-expert distinguish between different data structures? Is data structure delineated using the concepts: type of measurement, nominal, ordinal, interval and ratio scales or a discrete versus continuous classification; the number of characteristics measured (univariate, bivariate and multivariate); and, the nature and number of batches of data (related or independent, one two or more).
- Are there any other features that the professor-expert uses to define data structures of a more complex kind and hence determine the form of analysis?

The structure in this chapter is associated with categorical data (figure 3.10). The tools discussed provide different methods of displaying this type of data.

Figure 3.10 Exploring Categorical Data

Exploring Categorical Data

Data are measured by categories
Display frequences

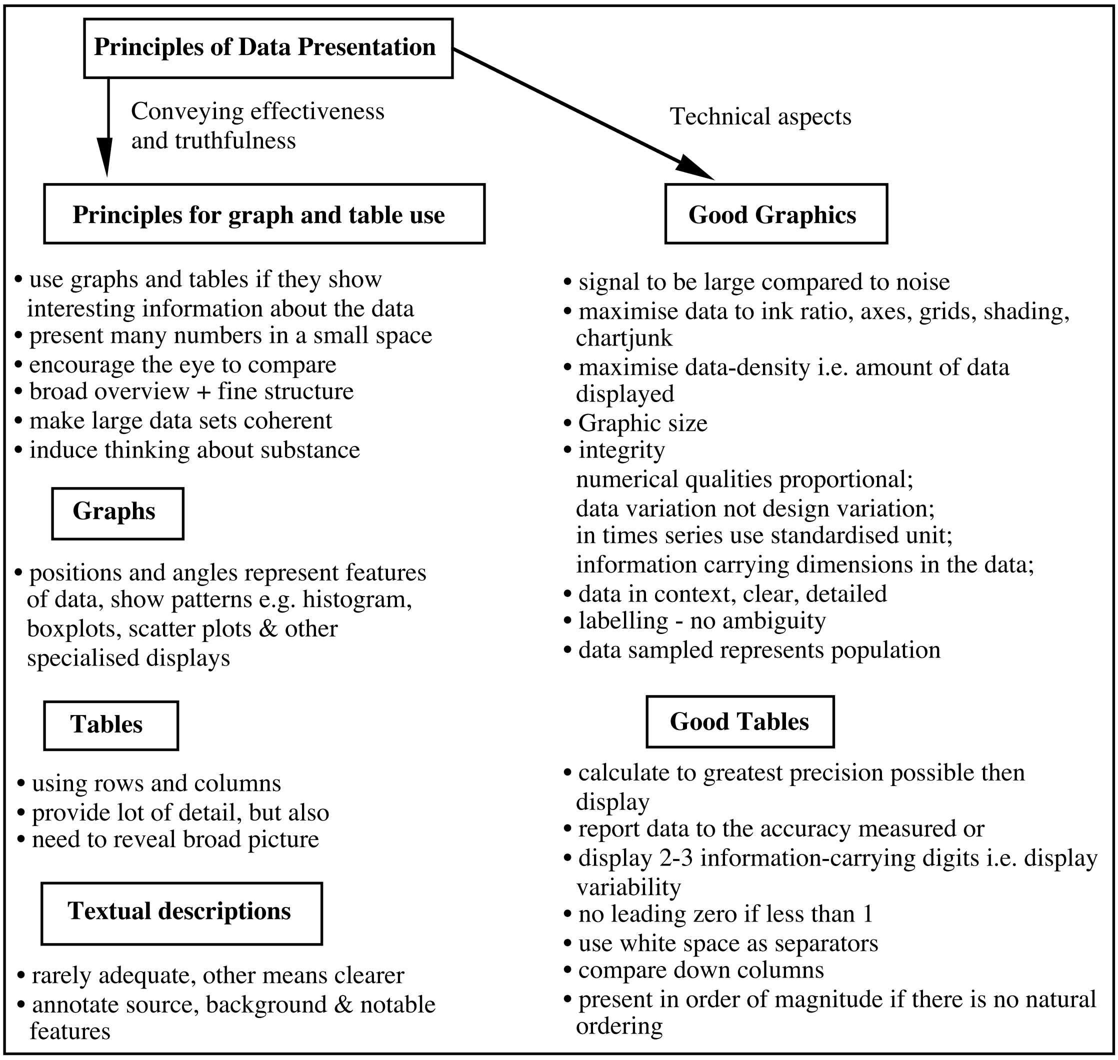
Structure	**Display principle**	**Technique**
One batch	Proportional to ... Some area of ...	Pie chart Bar chart
Two or more batches	Heights of bars in each batch proportional to number; total height standardized	Side by side bar charts Stacked bar charts
Pairs of categorical values		Three dimensional bar chart
Value pairs are numerical and categorical	Height corresponds to proportion Predicts proportions between 0 and 1.0	Stacked bar chart

The issue as to how to ascertain structure in data is examined more fully when the participant-researcher unpacks her expertise in Section 3.5.3.

Chapter 5 draft text: Principles of Data Presentation

The tools and techniques used to explore data are often the same tools used to communicate to others. For this reason, a chapter covering the principles of data presentation, Chapter 5 (Griffiths & Stirling, 1994), followed the exploration of data. The emphasis in the presentation of data rests upon the communication of information with clarity and truthfulness. To do this there are several principles of data presentation to observe when drawing tables, graphs and writing text, as mapped in Figure 3.11.

Figure 3.11 Principles of Data Presentation



In the chapter on effective presentation, reference is made to the display of data that represents the population. This leads to the next chapter that is on the appropriate collection of data. However, before this issue is examined the import of exploratory data analysis, as found in the text, is re-examined and confirmed in the light of observations of the professor-expert.

Exploratory data analysis: a commentary on chapters 1-4 through observation of the professor-expert

The emphasis in Chapters 1-4 on exploratory data analysis could be interpreted as the product of the statistical educator rather than the practising statistician. One of the professor-expert's research grant applications states that, 'education should reflect statistical practice' (Griffiths, Harper, Hedberg, Fasano, & Rayner, 1992). This comment regarding the relationship between statistical education and the practice of statistics suggests that the professor-expert would seek to incorporate those fundamental aspects of his work as a practicing statistician into the teaching of students.

Observation of the professor-expert whilst jointly engaged in the analysis of a project, *The role of stress in the health of the population* (1991-92), provided another source of evidence that the professor-expert places an emphasis on the exploration of data. The participant-researcher's notes, work records and recall of the experience with the professor-expert and other consultants in the years prior to, and during the undertaking of this thesis, suggested that when compared to other experts the professor-expert placed a much greater emphasis on exploratory data analysis. Confronted with the data from over fifty variables the professor-expert engaged in extensive exploratory data analysis. Frequency distributions were produced for each categorical variable, means and standard deviations for each continuous variable. Histograms were produced to show the distribution of male and female data for all continuous variables, and scatter plots were made of each pair of continuous variables whilst within sets of variables (for example, eleven questions on work-stress) cross tabulations of all pairings of categorical variables were made. Reviewing the print-outs, the professor-expert looked for what the data were saying:

- What was the typical value of each variable?

- Were there unusual data points (potential errors) that should be checked against the original source?
- Were there categories which should be combined (for example, combining categories *used to smoke* with *still smoke* to compare with *non smokers*)?
- Were the minimum and maximum values reasonable?
- Were there sufficient observations on each variable for the variable to be useful? (For example, for a dichotomous variable, are both outcomes common or is one so rare as to be of little value as an indicator?)

Extensive data exploration took place before formal regression and logistic regression analyses. The professor-expert appeared extremely patient when examining and exploring data.

In many ways, textbook analysis reveals little of the unique perceptions of the professor-expert. It was during the application of these techniques described in the text that the questions arose:

- What sense does the professor-expert make of these graphs and numerical summaries? and,
- What does the professor-expert see?

This same issue, which arose whilst observing the professor-expert examining output from *The role of stress in the health of the population* project, arose again when teaching students exploratory data analysis. Students can be taught to draw the various graphs and to calculate numerical summaries, but it is far more difficult to teach them to see what is in the picture and to identify and interpret what they are not seeing! The professor-expert can recognise not only the presence of desirable and undesirable features of data but the absence of these same features and make meaning from what is seen. By looking at the exploratory data analyses the professor-expert accepts what it says or looks beyond to further analyses. The expert sees more than that taught in the text when looking at data. For example (Griffiths et al., 1992) said: 'unless there are very obvious differences between batches, be careful about how strongly you proclaim these differences' (Chapter 1, p. 36). What is an obvious difference for one student may not be an obvious difference for another. Another example encountered in discussions with the professor-expert had to do with the shape of the tails of the distribution. The extent of the discussion in the draft text was as follows: 'Note that the discrepancy in the tails, although small, could be important' (Chapter 9, p. 9). Whilst a novice may well focus on the general shape of the distribution to indicate normality the expert is aware that the appropriate shape in the tails of distributions is of greater importance when testing hypotheses. Examining plots involves looking for the centre of data, minima, maxima, the shape of the distribution, the spread of the data, patterns, for the unusual and unexpected. The student needs to attend to all these features of data to draw meaning from what is observed. Looking to see what is both present and absent in data is more difficult for students than the construction of any given plot. Hawkins et al., (1992) made a similar statement in reference to the topic measurement:

> It is important to realise that it is this skill of judgement which must be taught, not least because a different judgement on the part of our students may lead to radically different ways of illustrating the data. (p. 44).

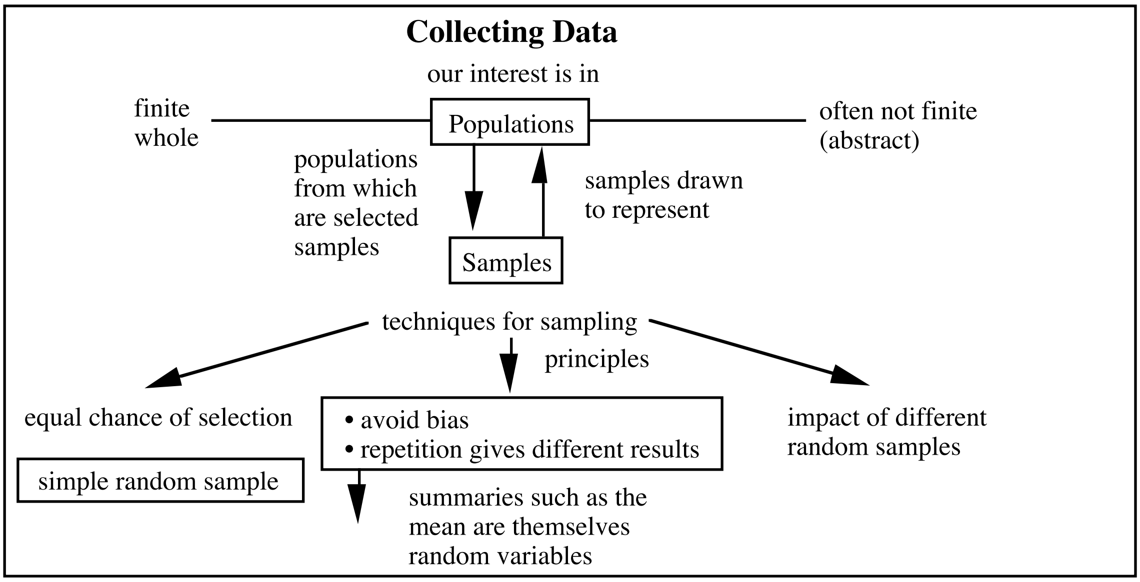
The comparison as to what one sees in data is not just between expert and student. The participant-researcher and other teaching assistants used model solutions, written by another expert (#2), to mark students' assignments. Questions often arose in viewing this material:

- Would all experts see the same things in data?
- How dense are data points before they are interpreted as being a cluster of points? and
- How great is the distance between data points before there is an interpretation that there is a gap in the data?

Chapter 6 draft text: Collecting Data

The chapter, *Collecting Data*, begins with the distinction between a finite population and a sample and leads students to an examination of how to draw a sample in order to represent the population of interest. The batch of data referred to, thus far, is defined as the 'whole population of interest' (Chapter 6, p. 1) whilst the sample is a group which is used to 'describe the distribution of this characteristic within a wider group of individuals' (Chapter 6, p. 1). This distinction between populations and samples and the subsequent principles of sampling are mapped in Figure 3.12.

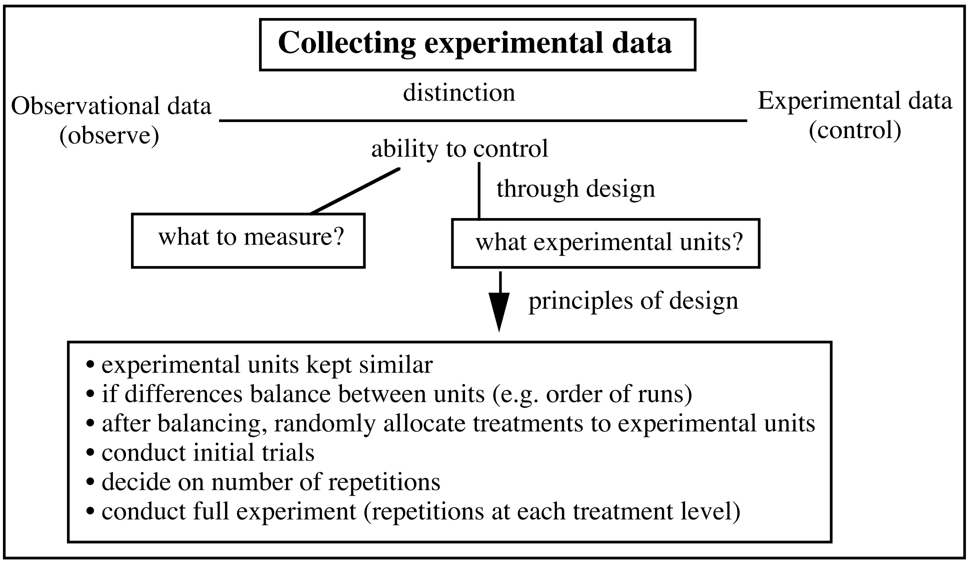
Figure 3.12 Distinguishing between samples and populations



The second major set of concepts has to do with the principles involved in collection of experimental data as distinct from observational data. These are mapped in Figure 3.13.

For this section on collecting data, reliance has been upon the text as neither concurrent nor prior experience involved the participant-researcher with the professor-expert in the design of experiments. In the text, the design of experiments formed a minor component, being but one of the twelve chapters, as was deemed appropriate for an introductory subject. (Indeed many Statistics texts have nothing on design issues.) This situation mirrors the participant-researcher's experience as consultant where most clients seek help at the analysis phase rather than the design phase when the first seeking of help is most appropriate.

Figure 3.13 Collection of Data

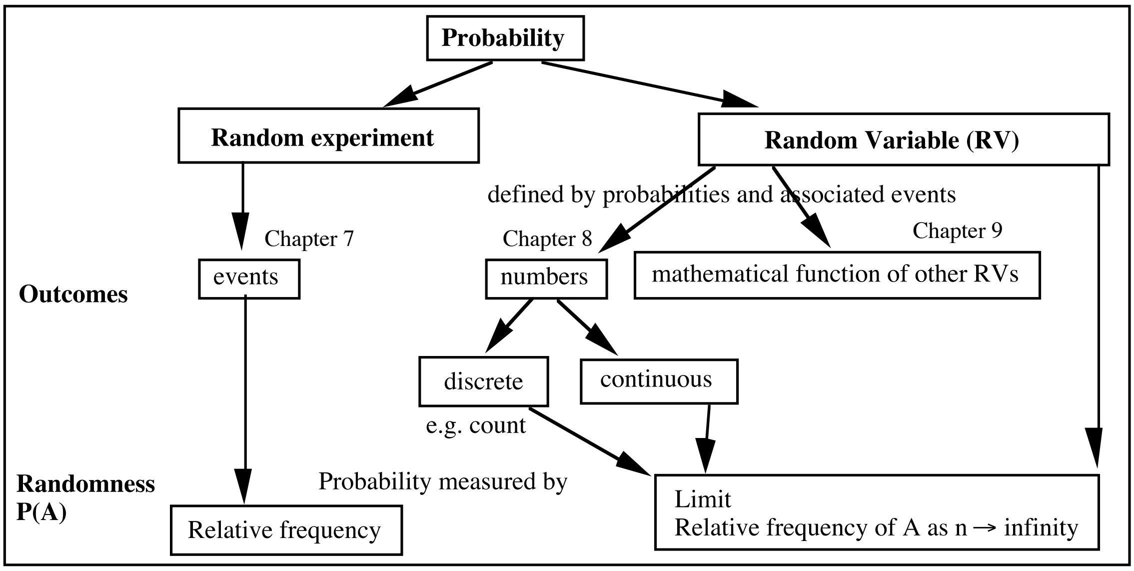


Probability and Random Variables: textual analysis (Chapters 7-9 draft text)

The unpacking of this topic involved mapping the text, observation of the professor-expert during the development of a pilot multimedia module on probability and the writing of a joint paper (with two other experts) called *The powers of two: How and why should we teach probability in introductory statistics courses?* (Micheal, Griffiths, Porter, & Stirling, 1993)

The global structure of the probability and random variables topic was drawn from three chapters. As can be seen in Figure 3.14, the subject matter in Chapter 7 was the probability of an event happening, for Chapter 8 it was random variables as defined by probabilities in conjunction with the random variable's numerical outcomes. In Chapter 9, the topic of random variables was extended to include functions of one or more random variables.

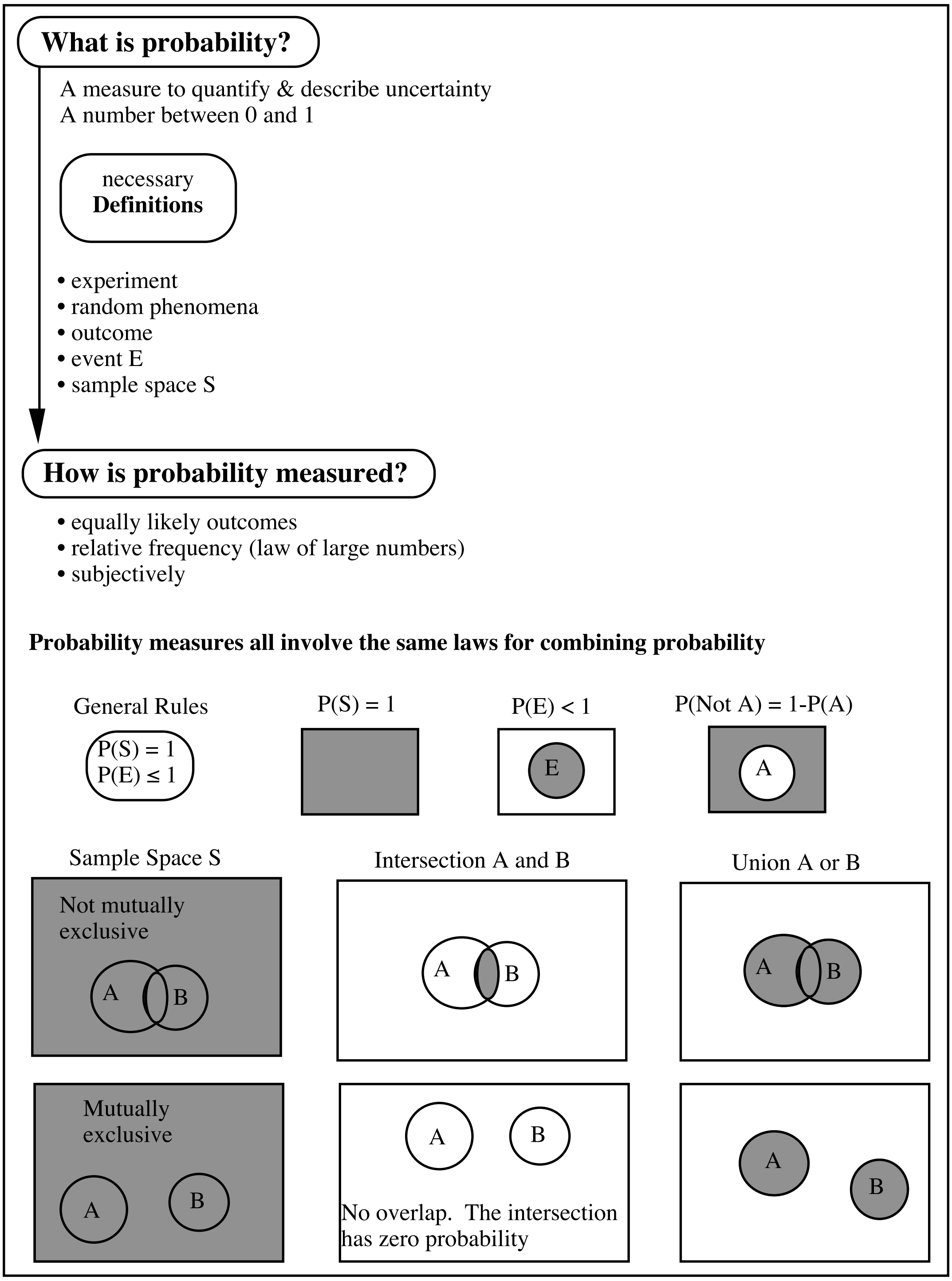
Figure 3.14 Broad structure of the probability topic



Chapter 7 draft text: Probability of an Event

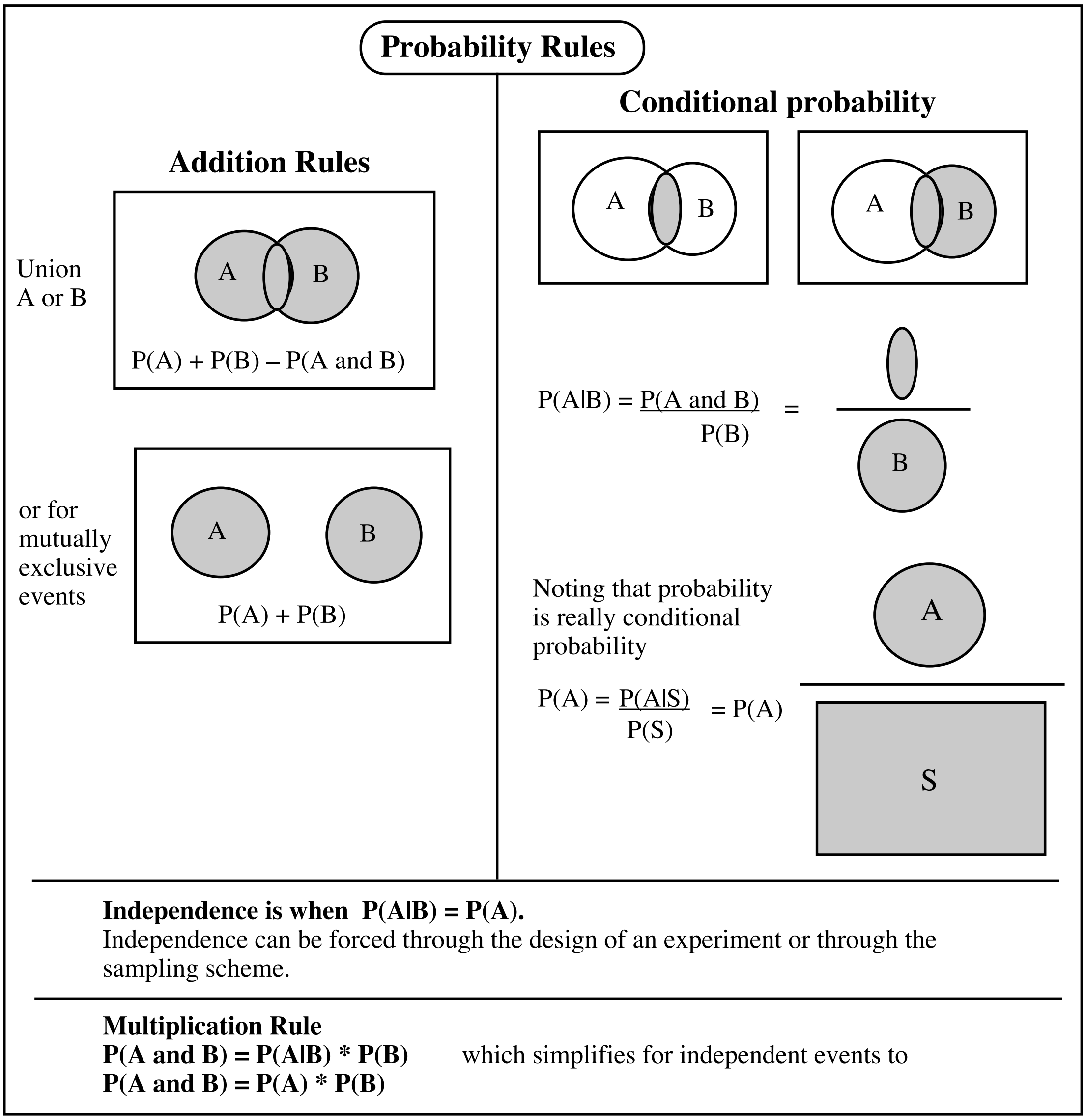
On this topic, the professor-expert emphasised the visual representation of problems, for example, tables or tree diagrams, in order to determine the probability of some event, rather than automatically using less intuitive combinatoric methods. This chapter is steeped in definitions of terms such as, experiments, random phenomena, outcomes, as the student is provided with a language to understand the topic 'What is probability?' (refer Figure 3.15).

Figure 3.15 Probability Concepts



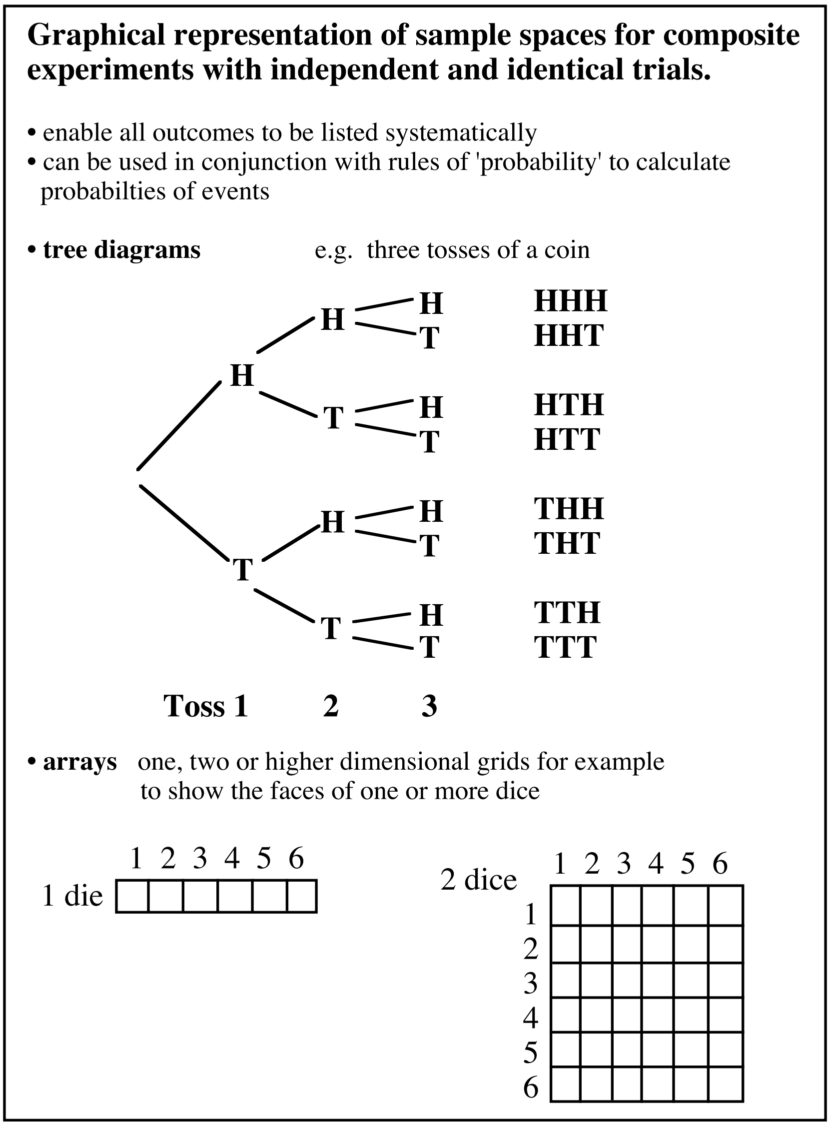
The chapter contains considerable material relating to the procedures or techniques regarding how to calculate probability (refer Figure 3.16).

Figure 3.16 Probability Rules



Students were also provided a visual means of representing sample spaces of probabilistic experiments (refer figure 3.17).

Figure 3.17 Visual representation of sample spaces



It was evident when designing multimedia materials and developing student laboratory tasks and assignments for use with the computer package *Models'N'Data* (Stirling, 1995) that the professor-expert sought ways of simulating probabilistic situations so that the learner could learn probabilistic reasoning through experience. A theoretical approach to teaching and learning probability was not deemed adequate. Students needed to develop their understanding through the observation of repeated experiments.

Many other ideas were covered in this chapter. They included, ideas associated with axiomatic probability, odds for and against an event happening, the use of tabular representations of data for the purposes of determining joint and conditional probabilities, an extension of the work on representing tabular data, intuitive means of reading probability from the tables, and, Bayes formula. These comprised what the professor-expert would consider basic probability concepts. The chapter also included the topic *Markov chains,* applicable where there is a sequence of dependent experiments. This topic was included both to extend and acquaint students with some interesting, yet relatively simple applications of the rules of probability.

Chapter 8 draft text: Random Variables - Modelling Numerical Data

In contrast to Chapter 7, where the randomness of a random experiment was described by the probabilities associated with all possible events in the experiment, this chapter considers a particular form of random experiment where all the possible outcomes are *numbers*. 'These numbers and the probabilities associated with them define a random variable' (Chapter 8, p. 1).

The initial discussion in the text provides examples of discrete and continuous random variables and supplies the notation that is used to distinguish the random variable (upper case letters, X, Y...) from outcomes (lower case letters, x, y...). An example is used to illustrate a simple random sample of n values and the relative frequency definition of probability that 'defines the probability for the outcomes' (Chapter 8, p. 3). The work on exploratory data analysis is recalled as it is revealed that 'numerical or graphical summaries based on a sample can be used to approximate (estimate) the corresponding summary for a random variable' (Chapter 8, p. 4). From a teaching perspective it is significant that the focus when examining the nature of random variables is again on familiar tools. This includes, for example, bar charts (or probability function as the limit is reached), histograms (or density functions as the limit is reached), medians, means, quartiles, percentiles, standard deviation, variance, and the inter-quartile range. Too often Statistics seems to students be a never-ending list of different things, where the meaning and usage is often not clear. Here students use familiar tools.

The Frames Type II method (West, Farmer, & Wolff, 1991) was used to represent the major concepts (defining characteristics, probability, sample mean, population mean, variance, and assessment of fit) for discrete random variables (refer to the earlier demonstration of a frame Type II, Table 2.2). The usefulness of the frame is that it provides for students the general form for calculating, for example, the mean or variance whilst at the same time indicating that there is an algebraic equivalent form which provides an easier method of calculation. In this manner, students are able to identify both the similarities and dissimilarities of the knowledge regarding the various distributions. The frame could, but for the limitation of table size, be extended to include the shape of distributions, the cumulative distribution functions for each random variable as defined in the text. It is these same concepts that are of interest when considering other discrete and continuous distributions. The same defining features, the probability of some event, shape, centre, spread and assessment of fit, are explored for continuous distributions and specifically for the rectangular, exponential and normal distributions.

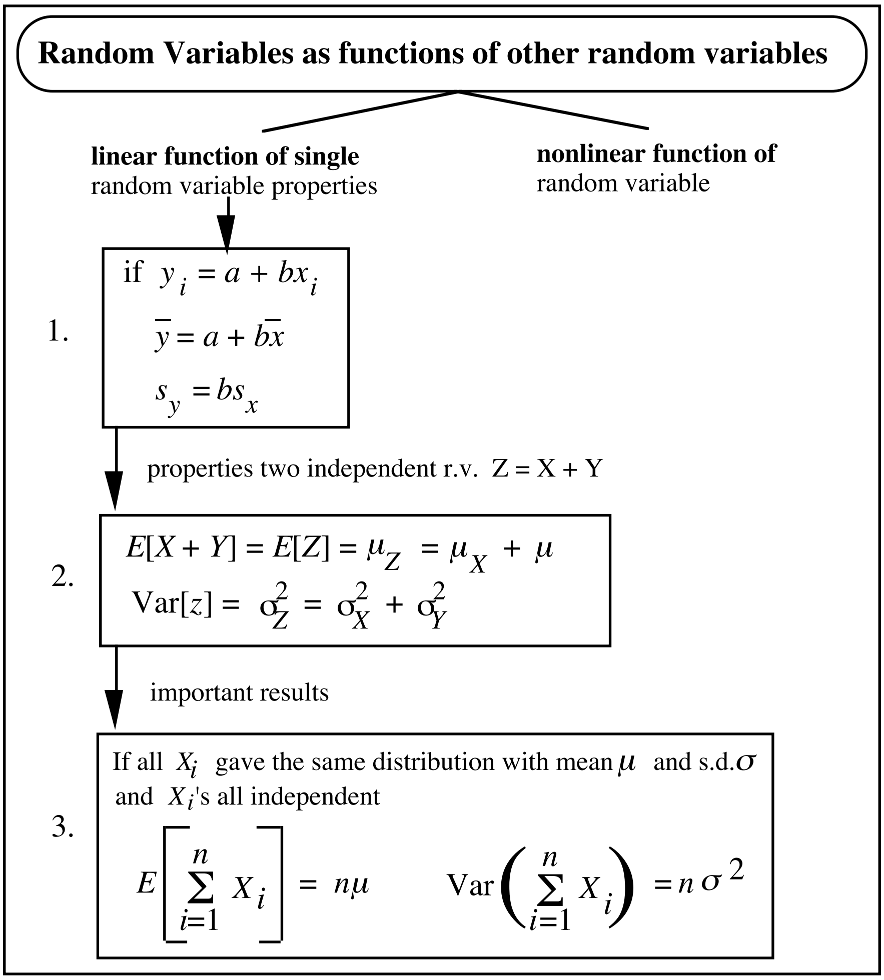
What has not been clearly shown in these frames but which is dealt with in the text is:

- the effect of the sample size increasing until the limiting form of the statistic is found;
- the shift from summation of the probability of outcomes when defining a discrete random variable to the use of integrals to find the area (and hence probability) of the random variable;
- point and interval estimates;
- the use of quantile-quantile plots for assessing the fit of continuous random variables; and,
- the incorporation of the number of parameters estimated into the determination of goodness of fit.

Chapter 9 draft text: Sampling Distributions

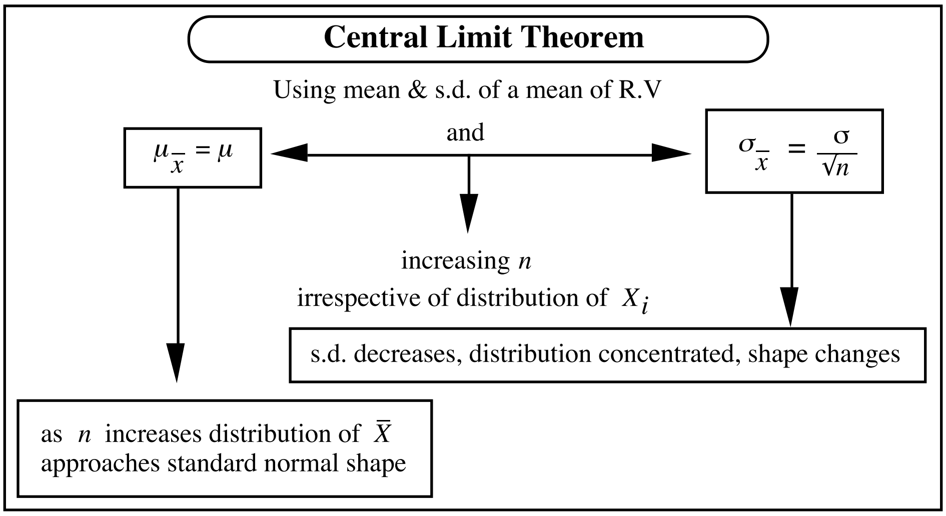
In this chapter, the case of a random variable as a function of other random variables is examined. As represented as in Figure 3.18, the mathematical properties of random variables are highlighted.

Figure 3.18 Functions of random variables



It is extremely difficult to convey the concepts associated with the central limit theorem, through textual material and for this reason, students' lectures are supplemented with laboratory simulations using a package called *Models'N'Data* (Stirling, 1995). With this package students are able to repeatedly generate experiments and build up the distributions of, for example, sample means in order to gain an appreciation of the next topic, the Central Limit theorem (refer Figure 3.19).

Figure 3.19 Central Limit theorem



Probability:A commentary on chapters 7-9 through observation of the professor-xxpert

As revealed by Chapter 0, *Why Statistics?* of the draft text, the professor-expert considered probability to be an integral component of the study of statistics. This perspective, referred to by some as Stochastics, (see for example Borovcnik 1990), is not adopted by all teachers of statistics. Hawkins et al (1992) describe how, for many educators, statistics is taught without probability or how it is taught as a separate module in a sequence of topics and thus, possibly, not as an integral component of statistics. In a later section the participant-researcher, whilst unpacking her own expertise, explores possible implications of learning statistics without probability.

The draft text was published five years later as, *Understanding Data* (Griffiths, Stirling, & Weldon, 1998). The text shows interesting changes. After the first five chapters, which could be called exploratory data analysis, the context of collecting the data (or design) is dealt with, followed by the strategies for dealing with uncontrolled variation (Statistical Inference). The chapters on *Probability* and *Random Variables* present in the draft (Griffiths & Stirling, 1994) are omitted from the published text (Griffiths et al., 1998) although essential basic material is incorporated elsewhere in the published text.

The changes in the text could be due to the influence of publishers or the additional author, or even to the maturation of the professor-expert. Between the draft and final publication it is also possible that no fundamental difference in expertise or the structure of expertise is discerned by the professor-expert. The text rationalises the omission of probability:

> This book contains little mention of probability. Probability modelling is a very useful basis for the description of situations involving uncertainty, and therefore underlies most advanced statistical techniques. However, students must be a little more comfortable with mathematics to appreciate probability models... It is important even for mathematically capable students to be immersed in data-based studies as portrayed here, so they do not come to believe that the probability models are more important than the information in the data modelled! Moreover, the role of models will be better understood if students are familiar with typical data based questions. (p. ix).

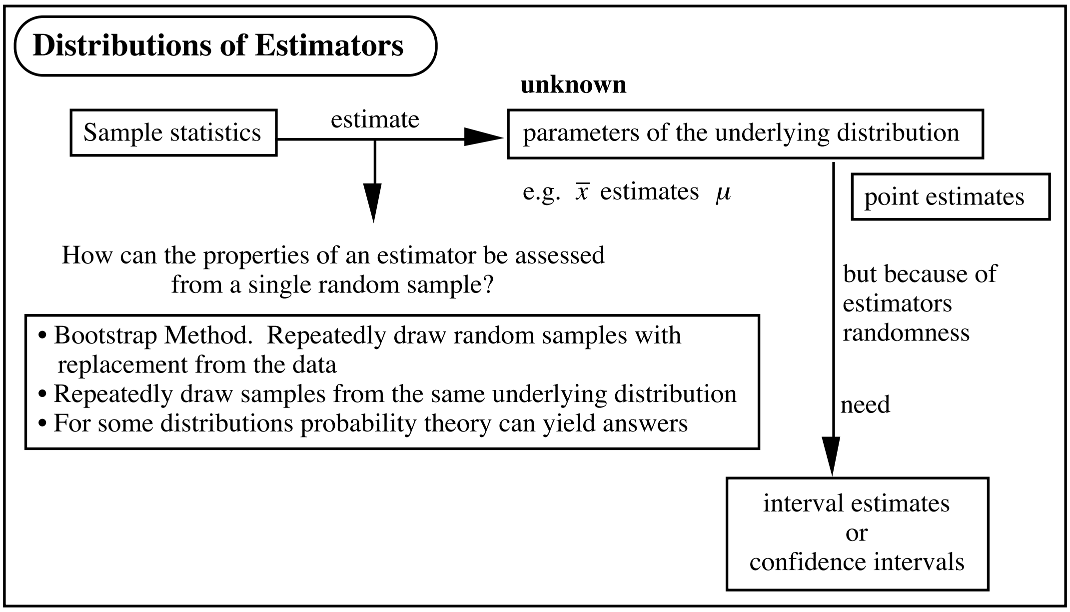
In teaching the introductory subject the year the text was finalised students were provided with the missing two chapters on probability! When questioned about the inclusion of probability the professor-expert initially indicated that, 'it provides a mathematical challenge that is necessary for mathematics students', and later added, 'but it is more than that, it provides concepts necessary for further studies.'

Statistical Inference (published text Chapter 7-8)

The chapters on Statistical Inference consider statistical inference for univariate data and inference for other data structures (comparing two groups; comparing three or more groups; models for linear relationships; and, bivariate distributions and correlation).

The initial sections of the chapter (refer Figure 3.20) introduce students to the idea of using sample statistics to estimate the unknown properties of the underlying distribution and the methods that may be used to approximate the distribution of the estimator. As the sample statistics are themselves random variables, the need for an interval estimate, as distinct from a point estimate, is also developed.

Figure 3.20 Estimation

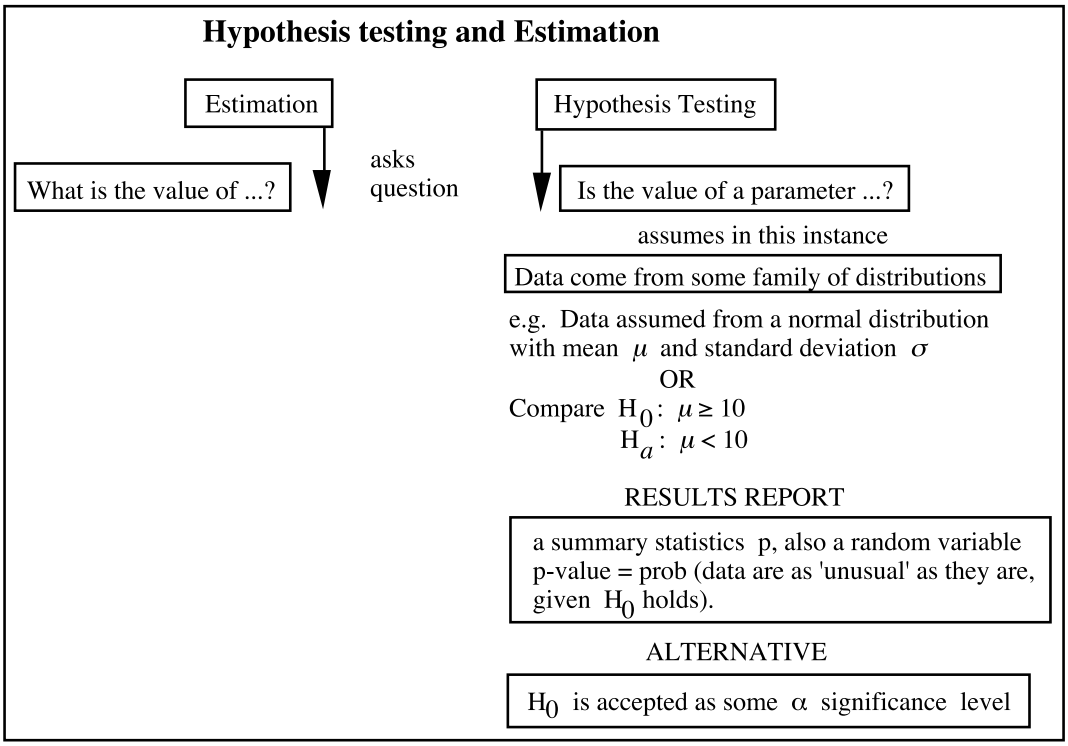


The distinction between estimation and hypothesis testing is then made (refer Figure 3.21). Estimation with point and interval estimates is to answer questions of the form 'What is the value of...?' (Chapter 10, p. 2), whilst hypothesis testing is concerned with providing input to a decision regarding questions such as 'Is the value of the parameter ...?' The hypothesis tests examined related some distribution, principally the Gaussian or t distributions.

Chapters in the draft text were incomplete. Hence, the unpacking was supplemented with a chapter, *From samples to populations - inference* (Griffiths et al., 1998). The initial distinction drawn in this chapter is between descriptive statistics, for which the data themselves are the object of the summary, and inferential statistics where the sample data are used to infer something about a population. Two aspects of inferential statistics were distinguished:

- estimation, describing a population; and,
- hypothesis testing which is evaluating claims about a population.

Figure 3.21 Estimation versus hypothesis testing



Throughout the text and draft text topics of exploratory data analysis, probability and random variables, and estimation three features of data and theory have been constantly examined. These were the location (for example, mean, median, mode), spread (for example, inter quartile range, variance, standard deviation), and distribution of data. In this chapter an example is provided for which knowledge about the whole distribution is more important than the mean or standard deviation. A similar example, used in my own teaching, would be where a shoe seller needs to know the distribution of feet sizes rather than the mean shoe size.

The distinction between point estimates, a single value, and interval estimates, a range of values within which the true but unknown value of the population mean is likely to exist, is introduced. To introduce the concept the authors state, 'the sample mean is a random quantity and varies from sample to sample, so there is likely to be inaccuracy in its use as an estimate of the population mean' (p. 304). The quantification of the confidence level about the accuracy of interval stated is then developed. The following ideas were used to develop a formal statement regarding confidence intervals (p. 307):

- If the interval is wide enough we should be very confident indeed.
- It needs to be narrow enough to be of practical use.
- The width of a confidence interval should in some way relate to the variability in the sample.
- The greater the variability in the sample the less information each value provides about the population mean, and hence the wider the confidence interval should be.
- As sample size, n, increases the sample mean...becomes more stable...the width of the confidence interval should therefore decrease as n increases.

It is statements such as these that the participant-researcher suggests aid in the development of meaning. Rather than a definition of what a confidence interval is, these statements provide the logic for what any such measure of confidence intervals should entail.

The theory is subsequently developed for the confidence intervals for:

- population proportions;
- differences between the means of two independent samples; and,
- differences between two proportions.

The professor-expert's views regarding the use of confidence intervals, instead of hypothesis testing, are in accord with Hawkins et al:

> To sum up: the confidence intervals, related to the particular value of the point estimate, are more informative than the results of significance tests. (p. 88).

The professor-expert proceeds, however, to distinguish between confidence intervals and hypothesis testing in addressing the general question, 'What does the sample tell us about the population parameter?' (p. 321). A confidence interval is used to answer, 'What parameter values are consistent with the data?' whilst an hypothesis test addresses the question, 'Are the data consistent with a particular value of the parameter?' (p. 322).

The final section introduces the notion of a p-value as a function of a random variable with attendant distribution properties which hold when the null hypothesis is true. The interpretation of *p-values* and usage of them as probabilities in relation to hypothesis testing ensues with the professor-expert data identifying the logic used as follows:

> The p-values for several other common hypothesis tests may be obtained in a similar way. They all follow the following logic.
>
> - Find an estimator of the parameter whose value is being tested
> - Find a formula for the standard deviation of the estimator
> - Evaluate the formula in step 2, assuming Ho holds (or at least find a good estimator for it)
> - Evaluate the z-score
> - z = (estimate of parameter)-(estimate of parameter under Ho)
>
>   standard deviation of the estimate.
> - Use the standard normal distribution to find how unusual such a z-score is; this is a p-value for the test. (Griffiths et al., 1998, p. 338).

For the professor-expert, each of these steps will call into play other knowledge, such as finding the standard deviation of the estimator. Students seeking help often have difficulty in recognising when one should select between the use of a standard deviation (s) and the use of the standard error ($s/\sqrt{n}$). The appropriate selection of the *standard deviation of the estimate* is not intuitive and is often the source of misconception for students. As Hawkins et al commented, in the context of a difference in random variables:

> ... students still believe that the variance operator acts like an expectation operator. Just as
>
> E(X-Y)=E(X) -E(Y)
>
> so students expect that
>
> Var(X-Y)= Var (X)-Var(Y) as distinct from Var(X) + Var (Y) (assuming that we are dealing with independent random variables, X and Y) (pp. 47-48).

3.3.2 The professor-expert: conclusions

The most succinct statement made by the professor-expert (Griffiths et al., 1992) characterised core statistical ideas as follows:

> A course in statistical literacy should have as its principal focus understanding variability. Most fields of human endeavour involve measurement. Without measurement, advances in those fields would be restricted and limited in value. Measurement involves variability and uncertainty. Statistics is concerned with understanding and interpreting variability. (p. 4).

For the professor-expert, statistics is about variability. The concepts developed throughout the draft text which further develop the domain of statistics, based on the chapters are the concepts: uncertainty, measurement, decisions based on data, data structures, exploratory data analysis, collecting data (design), probability and random variables, samples and populations, inference (hypothesis testing and estimation).

The explicated version of the professor-expert's text appears to place the emphasis on exploratory data analysis. Because of the many years of observation made of the professor-expert, the participant-researcher equates the exploration of data with the exploration of variation. However, this would not have always been the case. It was only late in the participant-researcher's study of statistics that there was an awareness that statistics was a study of variation. For example, when one speaks of a difference between means, this is one aspect of exploring variation; similarly looking for patterns in data, clusters, gaps, and outliers represent other ways of looking at variation in data.

The professor-expert utilises visual cues and looks at data, graphs it, summarises it, toys with it and only then, when all is explored, is there a move to formal analysis. In all the participant-researcher's contact with other professional statisticians or users of statistics (not a quantifiable statement) this characteristic has never been so dominant. The professor-expert examines intently the variation in data, making sense of it utilising data exploration tools. In making sense of data in this manner, the need to launch into theoretical or formal analysis is often lessened, although the principles of these undoubtedly guide the visual examination. The professor-expert in asking 'are you sure you want to test the value of a parameter, and not estimate it?' (Chapter 10, p. 3) also indicates his preparedness to estimate parameters, rather than test hypotheses, in the appropriate circumstances.

In the published version of the text the initial statement about statistics was changed. No longer do we have the statement: *Statistics is concerned with understanding and* *interpreting variability*. Rather, statistics is described in terms of data analysis in the real world context as follows:

> Statistics concerns the use of data to obtain information about real-life situations and problems. There are various statistical tools that help extract useful information from data. Techniques for summarising and displaying information are among the most important statistical tools, so we concentrate on looking at and summarising, the various common data structures. (Griffiths et al., 1998, p. ix).

Knowledge of the professor-expert suggests that the exploration of data is still very much an exploration of variation in data. This phenomenon of variability being subsumed under other terms became a major issue later in the thesis.

3.4 Other experts: episodes

To broaden the perspective on statistical expertise gained from the professor-expert an examination of others' statistical knowledge was pursued. The formal examination of others' expertise involved:

- chunking, concept mapping and framing of statistical materials for statistics courses, taught by other than the professor-expert;
- concept mapping other experts' statistical knowledge; and
- a review of published statements as to fundamental aspects of statistics.

Examination of subjects' texts and other materials

The professor-expert was chosen as someone who illuminated statistical theory. Students inspired the choice of text, and subjects to be examined. The students were those who sought help from the participant-researcher in the Learning Development Centre at the University of wollongong. In her role with students the participant-researcher used their notes and textbooks to ensure that the statistical language, symbols and formulae she used were the same as those provided in the respective courses (these vary between courses and between lecturers). On occasion, where multiple texts were set, discrepancies between them needed to be addressed (for example, use of equivalent but algebraically different formulae, or different symbols). The materials were also examined in terms of their content, formulae and structure, or lack thereof, as explanations needed to remain consistent with that which was taught. Explanations were generally expanded to provide students with some sense of data structure (type of questions asked, univariate or multivariate responses or explanatory variables, type of measurement and other features used, as discussed in Section 3.8.3) so that they could determine appropriate analyses. In this manner, the participant-researcher became aware of the contrast between the professor-expert's manner of exposition and that of other lecturers.

The participant-researcher drew upon the introductory textual material for three subjects. These included:

- statistical material written and taught by a statistician to a third-year level subject for Psychology students;
- an introductory statistics textbook for first-year Psychology students; and,
- the text used for a quantitative methods subject for Commerce students.

Concept mapping of other experts' knowledge

Approaching the task of identifying fundamental statistical concepts from another perspective the participant researcher drew on two concept mapping exercises with other experts. These were:

- a concept mapping exercise with staff and graduate students in the Department of Statistics at the University of Wollongong; and,
- a concept mapping exercise with groups of statisticians (including the expert) at the *STATISTICS '93* Conference at the University of Wollongong.

3.4.1 Other experts' course materials: Statistics for first-year psychology students

The text *Fundamental Statistics for the Behavioural Sciences* (Howell, 1995) is a typical reference selected for first-year psychology students undertaking their first course in statistics. The Preface addresses the question 'Why Statistics?' and, in accord with the professor-expert places the need for statistics in a real world context:

> Statistics is not really about numbers; it is about understanding our world... an important activity for statisticians is to answer such questions as whether cocaine taken in a novel context has more of an effect than cocaine taken in a familiar context. But let's not forget that what we are talking about here is drug addiction. (p. xv).

In the first chapter, multiple meanings of the term *statistics* are explored:

> …*statistics* refers to a set of procedures and rules for reducing large masses of data to manageable proportions and for allowing us to draw conclusions from those data...[The second meaning not used in this text is exemplified] by such statements as 'statistics show that the number of people applying for unemployment has fallen'... A third meaning of the term is in reference to the result of some arithmetic or algebraic manipulations applied to data. (p. 2).

In later chapters, statistics is characterised as two areas, descriptive statistics (including statements about averages and variability) and inferential statistics (making generalisations from samples about populations, taking into account the number of observations needed with greater variability). Probability, discussed later in one of twenty-one chapters, does not assume any great importance. The distinctions between statistics and parameters, random and non-random samples are drawn. Contrary to the professor-expert, Howell's emphasis is upon hypothesis testing, not the estimation of parameters which is described as 'a pretty dreary subject' (p. 7).

Howell continues in the introductory chapter referring to a decision tree which aids the choice of statistical procedures. Decisions are made with regard to:

- the type of data (measurement and categorical data - alternatively referred to as quantitative data and frequency or count data);
- questions (differences versus relationships);
- number of groups or the number of variables; and
- whether variables are independent or not.

Reference is made to scales of measurement that 'some textbooks' also use to categorise the ways of manipulating data.

The sense that statistics is about variability is not evident at this early stage. There is an emphasis on how to select techniques to address particular forms of questions. The text later includes a large number of statistical techniques:

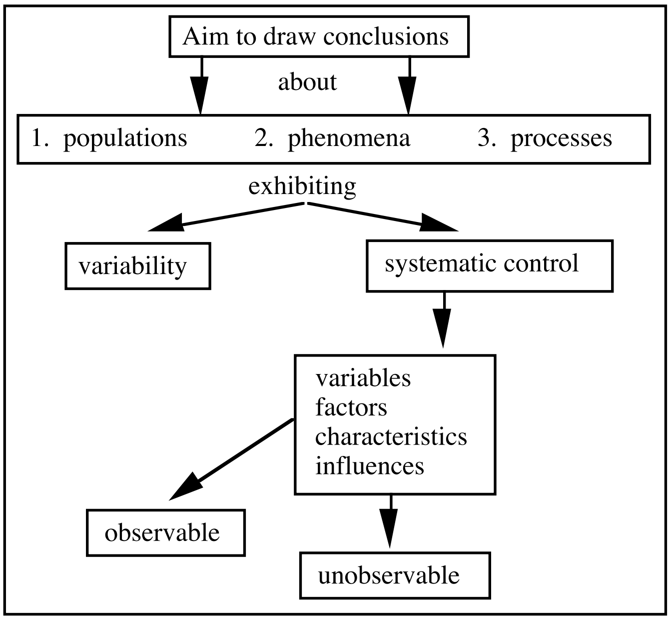
- correlation, regression, multiple regression;
- hypothesis testing for one, two related and two independent samples;
- analysis of variance, factorial analysis of variance, repeated measures analysis of variance;
- chi-square, non-parametric and distribution-free statistical tests.

In later phases of the thesis a more extensive analysis of textbooks is conducted to examine the extent to which the central role of variability is made explicit to students.

3.4.2 Other experts' course materials: Statistics for third-year psychology students

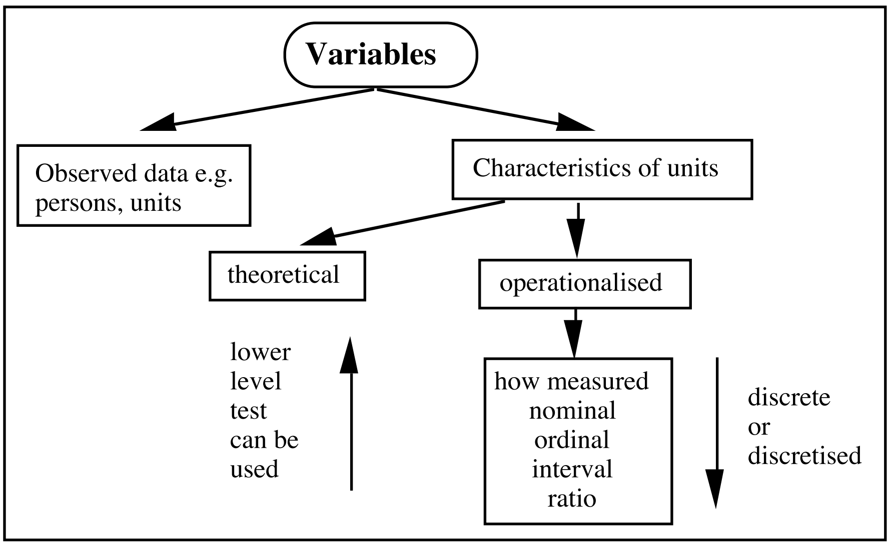
Written by a statistician, the student material examined was for third-year students in psychology. As it was semi-public material rather than formally published, the author (and teacher) and year has not been identified. The mapping (refer Figure 3.22) of the Overview chapter yields similarities and differences with the professor-expert. These similarities include addressing the need for answers to questions about populations, phenomena and processes that exhibit variability. Also included is the systematic control of variables, factors, characteristics and influences, some of which are observable but some of which are unobservable.

Figure 3.22 Overview of statistics



This expert defined data types for variables (refer Figure 3.23) and linked the data types to the tests that can be used in analysis. Variables were also classified according to whether or not they are dependent (language used by social scientists for response) or independent, with the further characterisation of independent variables as explanatory, stimulus, causative, factors or treatments.

Figure 3.23 Data Types

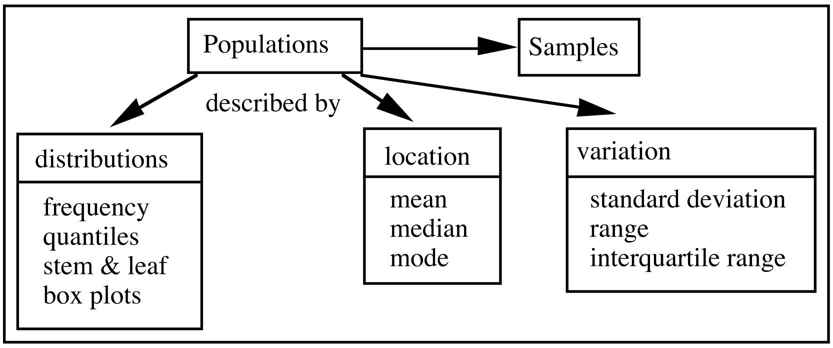


The analysis of data was deemed dependent upon three features of the data. These were the number of variables, independence and measurement scale. The professor-expert did not make this explicit link 'because multivariate analysis is essentially beyond the scope of 'his' book'.

According to this statistician (#2), variability is caused by, measurement error, genuine randomness and the influence of unmeasured factors. In a later description variability is described as occurring because of measurement, *real world processes* and sampling processes.

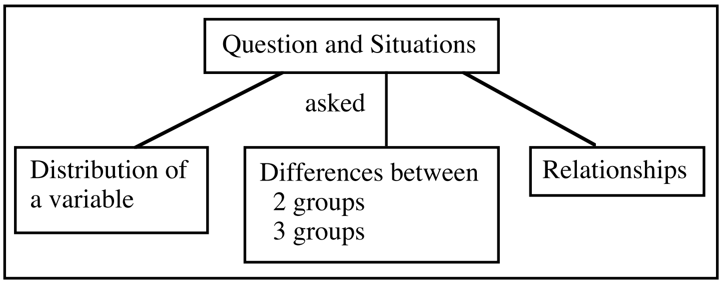
Students are introduced to the concepts of population and sample. As for the professor-expert the key features of data - location, variation and distribution - are commented upon (refer Figure 3.24).

Figure 3.24 Distribution, location and variation



The functions of statistics in terms of point estimation, interval estimation and hypothesis testing were described. This was followed by the issue of the nature of questions asked and situations examined (refer Figure 3.25). These themes in relation to statistics were addressed more comprehensively by Speed (1986, refer, section 3.4.6) but less explicitly by the professor-expert.

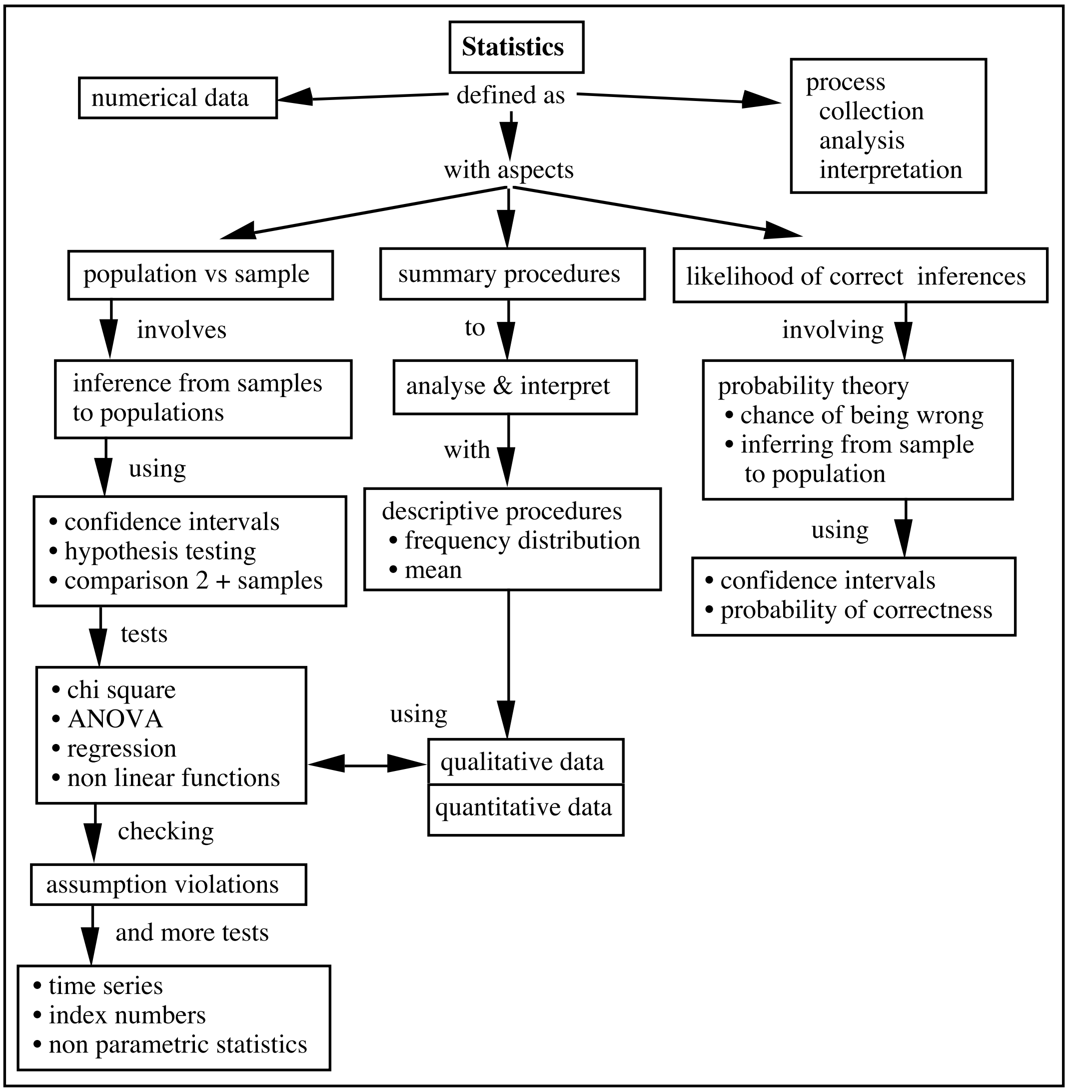
Figure 3.25 Questions asked in statistics



3.4.3 Other experts' course materials: Statistics for first-year commerce students

The text *Statistics for Business and Economics* (Lewis, O'Brien, & Thampapillai, 1990), used by first-year commerce students, provided a contrast to the work of the other experts. The context for the text is that organisations need to *process and consider information* in order to make decisions. The Overview Chapter (refer Figure 3.26) includes many advanced analysis techniques. Exploration (summary procedures), inference (from samples to populations using hypothesis testing and confidence intervals, omitting estimation) and probability, in addition to the process of collecting data are included in this text and referred to in this chapter. Variability is not suggested as the *raison d'etre* of statistics.

Figure 3.26 Overview chapter for Commerce students



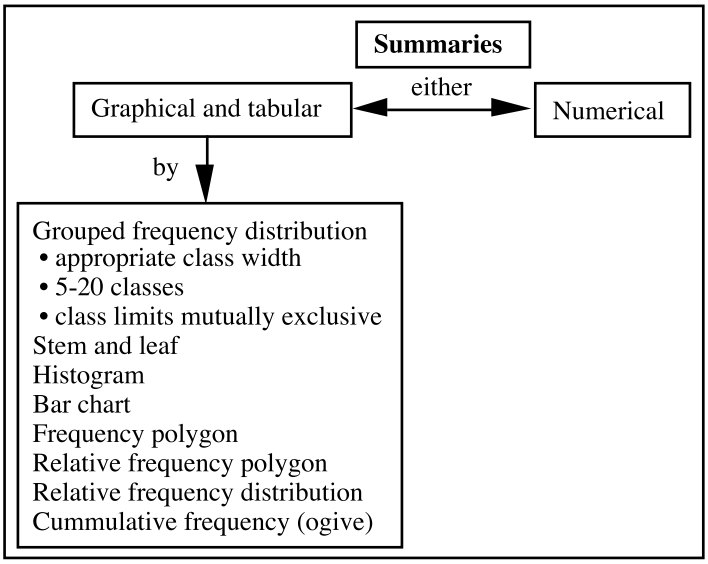
The second chapter distinguished between graphical and numerical summaries and then focussed on graphical summaries as in Figure 3.27. The material in these two chapters could be described as lexically dense. There was reference to terms that are laden with meaning for which there was no explanation. For example, following the early emphasis on summarising information, the introduction to the section on descriptive statistics is as follows:

> In Chapter 2, we dealt with methods of grouping data into frequency distributions. These can provide useful ways of summarising data. However, comparisons often need to be made among different sets of data and the use of frequency distributions for purpose of comparisons may prove to be arduous. An alternative method is to summarise the data in terms of **parameters** or **statistics**. It is often useful to calculate numerical values which summarise a data set with respect to its **central location** and **dispersion**. When data is from

> a population, the summary measures are called **population parameters**. When the data set is a sample, the summary measures are called **sample statistics**. (p. 19). (Bolding added).

The linkages between concepts were negligible, the nature of the material definitional (or declarative). The structure provided by the professor-expert of defining data types, for which particular graphics are more or less appropriate, was absent.

Figure 3.27 Graphical and tabular summaries



Whilst the selection of texts was opportunistic, they were typical of the texts that students bring when seeking help in these disciplines. The first subject taught by a statistician to health and behavioural students provides an overview, not unlike the professor-expert, with variability being of central concern. The other two, statistics subjects, taught by non-statisticians to psychology and commerce students respectively, utilised texts that did not include a focus on variation. Rather the texts were oriented toward techniques, that is, the *how to* of statistics.

3.4.4 Other experts' concept mapping: staff and students

Early in the development of the thesis, honours and postgraduate students and staff attending a departmental seminar were asked to 'jot down the six most important statistical concepts and then to add lines to indicate the connections between them'. The object of the exercise was to introduce the intent of this thesis, the unpacking of expertise. A two minute exercise, inadequate for the task, yielded some interesting differences between staff (refer Table 3.2) and students (refer Table 3.3) and within both groups. It heralded the shift from unpacking concepts (a search for the ideal and true structure of statistical knowledge) to recognition of different perspectives or constructions of statistical knowledge.

Three of the five members of staff, including the case expert, listed variability as the first of their central concepts; a fourth included it in the list, and a fifth, who focussed on topics taught to students, omitted variability as a concept. To glean more from the listing requires some interpretation of meaning assigned to the terms. For example, distribution-related concepts (Staff 5), distributions (professor-expert), probability (Staff 3) and random variables (Staff 4), could possibly all be interpreted as probability, and thus four of the five included this as a key concept.

Table 3.2 Basic Statistical Concepts - Other Statisticians

Staff	Concept 1	Concept 2	Concept 3	Other Concepts
Professor Expert	Variability	Populations Samples	Data Structures	Distributions Inference
Staff 1	Variability	Populations Samples	Multi factors	Inference Probability
Staff 2	Variability	Data Analysis	Decision Making	Design of experiments
Staff 3	Normal Distribution	Location Measures	Variance	Random Variables Linear Regression
Staff 4	Normality & other distribution related concepts	Graphical Presentation	ANOVA	Regression Multivariate Distribution-free methods

As with staff, the student conceptions (refer Table 3.3) indicate a variety of concepts being identified. Most interesting, and in contrast to the staff, is the absence of the concept of variability.

Table 3.3 Basic Statistical Concepts - Students

Student	Concept 1	Concept 2	Concept 3
Student 1	Collect relevant data that help make good interpretations	Good presentation of data	Enough testing
Student 2	Probability	Simulation of natural occurrences	Expectation, Randomness
Student 3	Analysing data to lead to important conclusions	Turn numbers into something meaningful	
Student 4	To be able to test models (distributions) to real world applications	Forecast and predict from current models	

This exercise was sufficiently revealing about the differences between statistician and student concepts, that a formal workshop at the Statistics '93 conference was conducted in order to provide further insight into what concepts are considered fundamental to the discipline of statistics.

3.4.5 Other experts' concept mapping: Conference participants

At the STATISTICS '93 conference a workshop was held for participants in order to explore their central statistical concepts. It was not possible in this situation to identify the participants as statisticians, statistics educators or *statisticians* from other disciplines. It was evident that there was a mixture of all three types of participants.

Participants were provided with written instructions (refer Appendix 3.1) and were asked to identify ten basic statistical concepts, and to subsequently link them in a concept map. Approximately 10 minutes was allowed for the task. Participants were then combined to form groups of four or five. The grouping of participants was accomplished through individuals teaming with others with the same coloured instruction sheets. The groups were then asked to develop a list of basic concepts and to map these hierarchically, identifying the links between concepts. Approximately 35 minutes was given to this task. Participants, should they be prepared to do so, handed in their individual maps, so that the group maps could be compared and contrasted to the individual maps. Four group maps, based on the work of 18 participants, were submitted. Thirteen individual lists of concepts and/or maps were collected. The professor-expert was a participant in this workshop, although his concepts/maps and group membership were not identifiable.

The concepts utilised in the individual list and group concept map for each of the four groups were compared to the reference set provided by the earlier analysis of the professor-expert. The primary concept in this reference set was variability, the others were: uncertainty, measurement, data structures, exploratory data analysis, collecting data (design), probability and random variables, samples and populations, inference (hypothesis testing and estimation) and real life problems.

Due to limited space, only a selection of the maps has been included. The first group map, was the only one which did not show variability as a major concept. It was included together with the accompanying individual maps that led to the composition of the group map. These revealed how the individual maps were transformed to a group map when teamwork was involved. The three remaining group maps that are presented all include the concept variability. The associated individual maps have been omitted.

Group 1

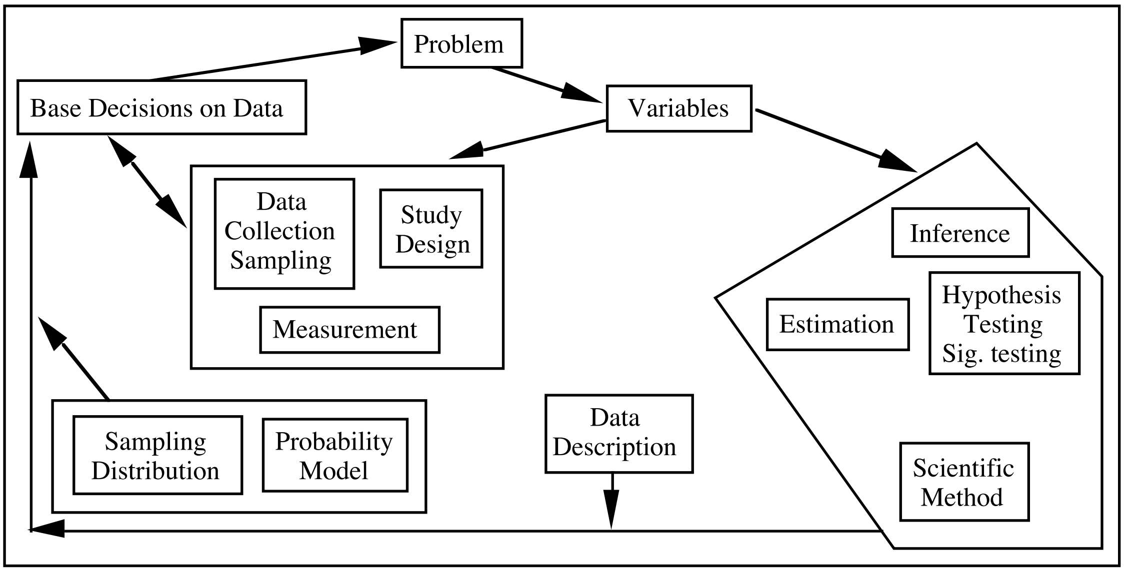
The group concept map, Figure 3.28 is based on the work of four participants (Figures 3.29 - 3.32). Interestingly, the group map does not include variation or variability as a central concept even though it is listed by three of the four participants. Included in the map were measurement, decisions based on data, exploratory data analysis, collecting data (design), probability[2] and random variables, samples and populations, inference (hypothesis testing and estimation) and real-life problems. The map contains most concepts considered central to the professor-expert's view. There was more similarity than dissimilarity with the professor-expert at the subordinate level of concepts. Two concepts, uncertainty and data structures were not included, although *probability model* will necessarily require some reference to both. Two additional concepts, *variables* and

---

[2]The terms used may differ slightly in meaning, for example, probability model may or may not have included random variables.

scientific method, were included. However it is difficult to know if a distinction was intended between scientific method and the cluster of concepts, *hypothesis, research design, data collection, study design* and *measurement,* which are often incorporated in descriptions of scientific method (Dolphin, 1992; Dye, 1996; Miller, No date). If it can be presumed that this is the structure through which one studies variation (that is, the variation is implicit rather than explicit) the perspective is very much in accord with the professor-expert. If this cannot be presumed, then the very essence of what statistics is has been omitted.

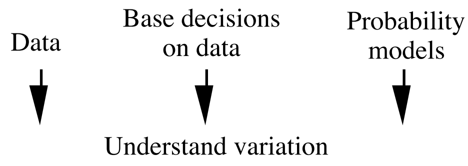
Figure 3.28 Group 1 Concept map



All the first participant's concepts (Figure 3.29) were included in the professor-expert's list. In making this statement it was recognised that this participant made a distinction between scientific method and notions of design and sampling; and the term data description has been equated with data exploration. The fundamental objective, as with the professor-expert, was to understand variation.

Figure 3.29 Participant 1, Group 1

1. Scientific method
2. Understanding variation
3. Probability model
4. Random variable
5. Sampling distribution
6. Good design of study leads to better analysis/study design
7. Inference (from sample to population)
8. Base decision on data
9. Significance testing/estimation
10. Data description
11. Measurements



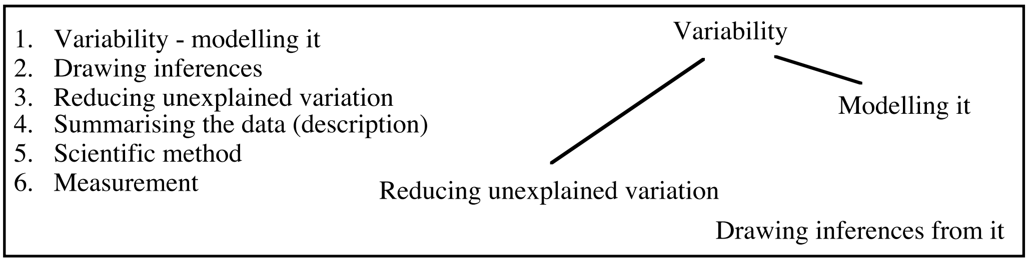
The second participant's list (refer Figure 3.30) was not as developed as the professor-expert's list. The rhetorical nature of the list brings to the fore the question of meaning of each term. In accord with Kelly's (1955) *Personal Construct Theory* it is quite possible that different meaning is attributable to the same terms, and for that matter, the same meaning might be attributable to different terms. For example, *looking at data sets* has been interpreted as being comparable to the professor-expert's *exploratory data analysis*, although with four chapters devoted to this topic there is much scope for there to be differences in meaning; that is, the subordinate concepts or links between them may differ. Issues such as '*what is confidence?*' can be considered to be included within the professor-expert's examination of the inference topic, but at a more subordinate level. On the surface there is a similarity between the reference view and this participant's, although it is not as developed.

Figure 3.30 Participant 2, Group 1

1. What is a sample?
2. What is a random process?
3. What can you gain by looking at data sets?
4. What is 'confidence'?
5. What is variability?
6. What is a model?

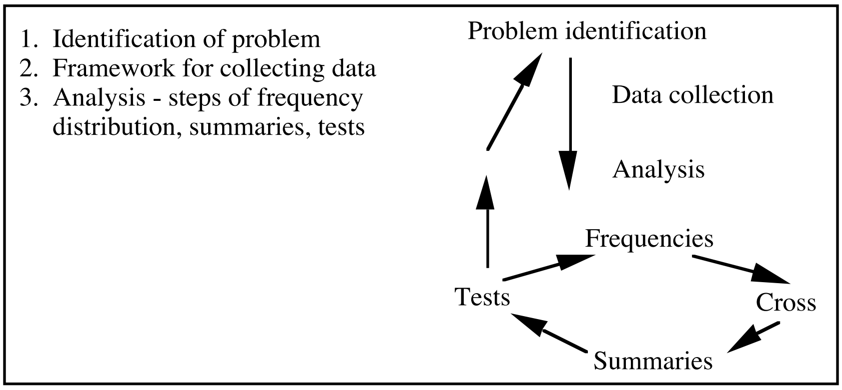
As with the first participant, the study of statistics for the third participant (Figure 3.31), was centred upon the concept of variability. The supporting concepts were all included in the reference list. There are some omissions, which may be due to the time allowed or to the conceptualisation of the participant.

Figure 3.31 Participant 3, Group 1



The fourth participant (Figure 3.32), provided a framework for the conduct of statistical work. There was a rudimentary process identified that began with a problem and that involved data collection and analysis. It could be taken to represent a basic *how to do it* guide to statistics. It omits the essence, that statistics is a study of the variation in data.

Figure 3.32 Participant 4, Group 1



Through the inclusion of the central concept, variability, and the inclusion of many of the subordinate concepts, the first three participants may be considered to have similar views to the professor-expert. There is, of course, much scope for the differentiation of views at the subordinate levels and in the linkages between concepts. There were insufficient data to examine this.

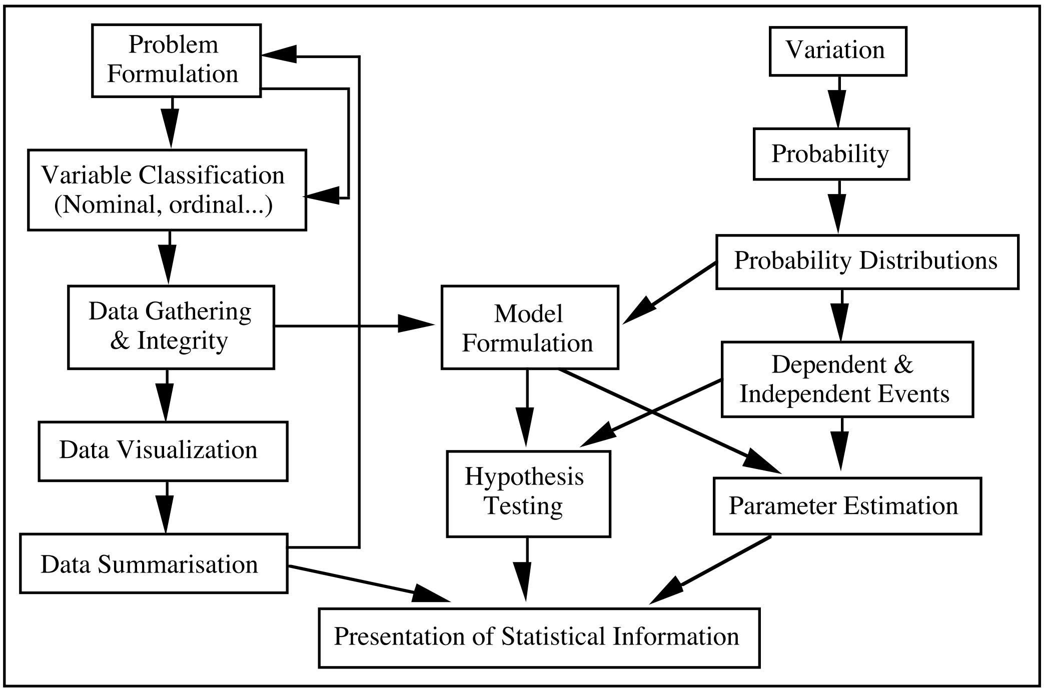
Group 2

The group concept map (Figure 3.33) was based on the work of four participants, although only three submitted their individual maps. This group map included variation, as a central or higher order concept, and variability in some form was listed by each of the three participants. All the concepts mapped could be considered as included on the reference list, although specified at different levels of inclusiveness; for example, exploratory data analysis is inclusive of concepts, such as data visualisation and data summarisation.

This group map was interesting from an educational perspective. The list of concepts on the left, headed with problem formulation, in one sense may be interpreted as a *how to do it,* whilst the concepts on the right, headed variation, reflects *the why*, the necessary theoretical underpinning for understanding what is done. In a crude sense it reflects the Gowin's Vee conceptualisation of good knowledge. This feature was also evident in the third group map.

Two of the three participants in this group had variability as a central concept. Whilst the group concept map was inclusive of the concepts generated by the professor-expert and had variability as a superordinate concept, none of the individual lists was comprehensive enough to be considered close to the reference list.

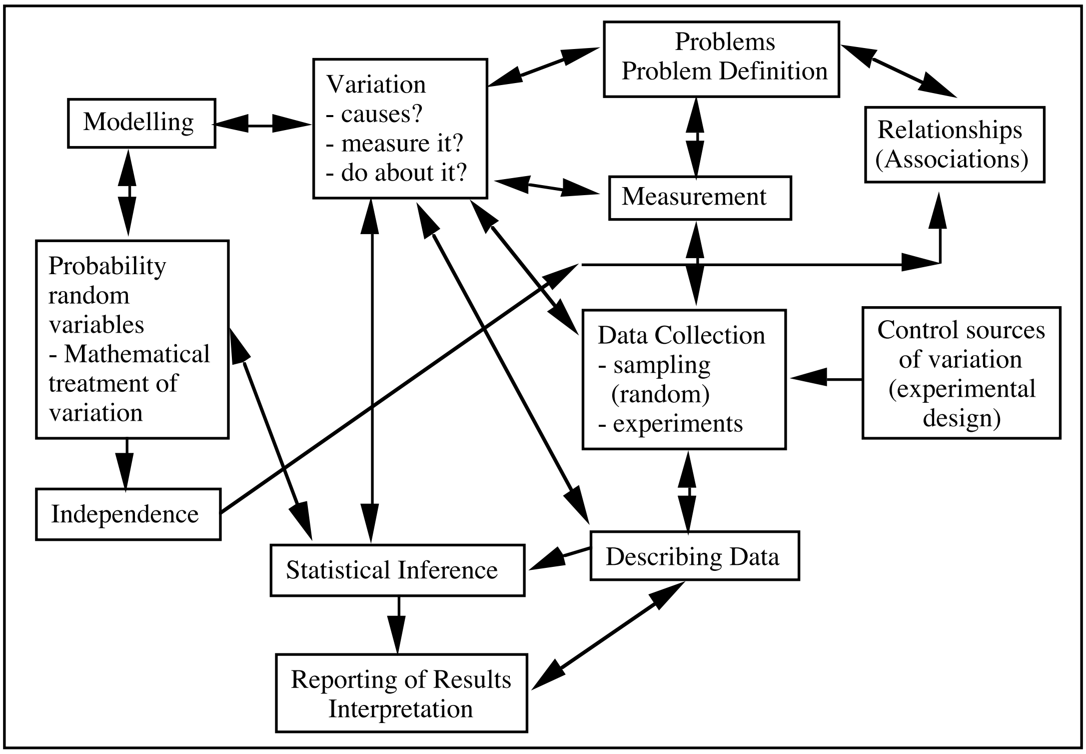
Figure 3.33 Group 2 Concept map



Group 3

The group map (refer Figure 3.34) was the work of five participants although only one participant submitted their individual effort. The group map clearly showed the centrality of the concept of variation. As for the previous group map the left hand side of this map showed the theoretical aspects, the *why of statistics*, whilst the other side was more akin to the statistical process or *doing it*.

Figure 3.34 Group 3 Concept map



Individual concept maps for the first three individuals (refer Figure 3.35) in this group had variability as a central concept, whilst one listed uncertainty (the first to do so). The first participant omitted reference to several major concepts (problems, exploratory data analysis, data structures, collecting data), whilst the lists for the second and third participant were poorly developed.

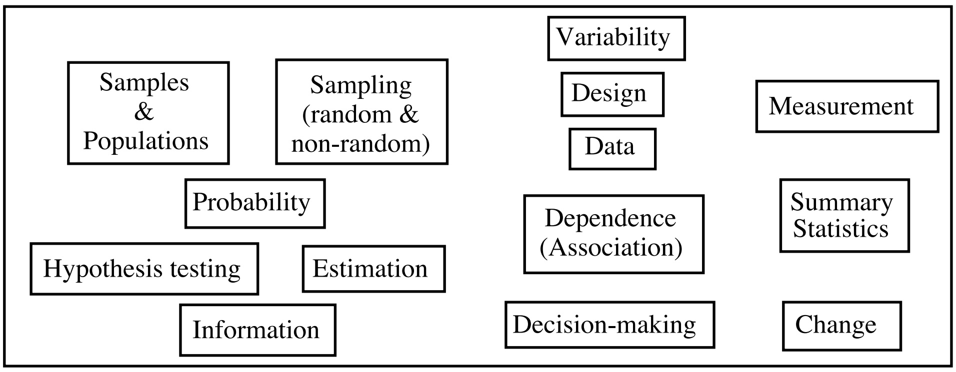
Figure 3.35 Participants 1-3, Group 3

Participant 1	Participant 2	Participant 3
1. Distributions 2. Variability 3. Random sampling 4. Random allocation 5. Variables 6. Measurement scales 7. Probability (p Values) 8. H0/H1 9. Samples 10. Populations	1. Variability 2. Dependence 3. Sampling 4. 'Change' 5. Estimation	1. Random Sampling 2. Samples & populations 3. Degree of uncertainty 4. Variability

Group 4

The fourth group map (Figure 3.36) was based on the work of five individuals, all of whom submitted their individual list of concepts. Whilst three of the five individuals in this group had variability as a central concept, the maps for all the participants in this group were poorly developed, with only one beginning to show the linkages between ideas.

Figure 3.36 Group 4 Concept map



As reference to the group map reveals, this group needed more time to link the concepts and possibly even to place appropriately the concepts in relation to each other. However, according to its position on the map, variability was the concept identified as superordinate. All the major concepts derived from the professor-expert's unpacking may be considered to be included on the map, with the exception of data collection, although it was not clear whether or not *data* was meant to encompass this process. Change seemed a little out of context and the inclusiveness of the term was not clear.

Conclusion: Other experts' Statistics '93 concept mapping

Whilst the time was limited by a 45 minute conference timeslot, it was considered that practitioners would in a short space of time be able to give those concepts which were of primary importance to them. A more considered and finalised version if obtainable, as with the professor-expert, might be many years in development.

It is interesting to examine the list of concepts, and the mapping of those concepts, and to recognise that the team effort appeared to be more than the sum of the individual efforts. Whilst at night, and after a long day of listening to conference papers, it was still evident that:

1) practitioners and statisticians found it difficult to articulate the basic concepts or it appeared that the concepts were not readily expressible;
2) even fewer linked concepts, although it was recognised that little time was given and that mapping ideas may have been a new idea to some participants;
3) the team approach to mapping engendered much discussion and debate and more comprehensive mapping of concepts;
4) all group maps included the concepts deemed important by the professor-expert;
5) ten of the thirteen individuals listed variability on their individual maps; and,
6) at the individual level there was much greater variation in the comprehensiveness of concepts maps than at the group level.

3.4.6 Other Experts: Perspectives on Statistics

Throughout the iterative process of reading the literature, unpacking statistical expertise and implementing the curriculum, different holistic perspectives emerged regarding the nature of statistics. These perspectives, some of which have already been expounded are described in this section.

Variability is the basis of statistical thinking

Various authors have endorsed the perspective that statistics is a study of variability (Cobb & Moore, 1997; Hogg, 1991; Zayac, 1990). Reporting on the *Second Australian*

Conference on Industrial Statistics, Robinson (1997) summarised Peterson's definition of statistical thinking as follows:

- seeing structure in randomness,
- seeing structure in variation: a deterministic and random component,
- knowing the consequences and importance of variability,
- knowing appropriate actions in response to variation,
- being able to identify causes of variation, and
- knowing how to avoid pitfalls due to causes of variability.

(Peterson, cited Robinson, 1997, pp. 10-11).

Returning to view the work of the man considered to be the father of modern statistics we find the following statement regarding statistics as a study of variation:

> The populations which are the object of statistical study always display variation in one or more respects. To speak of statistics as the study of variation also serves to emphasize the contrast between the aims of modern statistics and those of their predecessors. For until comparatively recent times, the vast majority of workers in this field appear to have no other aim than to ascertain aggregate, or average, values. The variation itself was not an object of study, but was viewed instead as a troublesome circumstance which detracted from the value of the average. The error curve of the mean of a normal sample has been familiar for a century, but that of the standard deviation was the object of researchers up to 1915. Yet, from a modern point of view, the study of the causes of the variation of any variable phenomenon, from the yield of wheat to the intellect of people, should be begun by examination and measurement of the variation which presents itself. (Fisher, 1958, p.3).

Variability and Uncertainty

In accord with the professor-expert, others provide support for the idea that the fundamental nature of statistics, arises out of a need to measure as well as to understand variability and uncertainty:

> Statistics is concerned with understanding the real world through the information we derive from classification and measurement. Its distinctive characteristic is that it deals with variability and uncertainty which is everywhere.(Bartholomew, 1995, pp. 5-6).

In another description Bartholomew describes variability and uncertainty as 'two sides of the same coin' (p. 3). Alternatively we find the commentary provided by Moore:

> What we offer beginners should reflect the conviction that statistical ideas are the key to sound thinking about data, variation, and uncertainty in all areas of life and work. (Moore, 1999, p. 137).

Statistics is a process

Wild (1994) argues for a 'wider view of statistics' and that 'we should consciously decide to make statistics the study of the process of scientific enquires' (p. 165). Wild adopts the definition, 'Statistics is concerned with finding out about the real world by collecting, and then making sense of data' (p. 164) in combination with a flow chart of the investigative process to define statistics. The investigative process includes the phases: real problems, curiosity; questions about the world; designing the method of data collection; collecting data; summary and analysis of data; and answers to original questions.

Probability is the basis of statistics.

> My thesis is simply stated: The core concept, around which all statistics should be based is probability. (Lindley, 1990, p. 33).

This assertion that probability is the core is often evident in the literature. However there are difficulties as the quote anecdotally attributed to Bertrand Russell reveals.

> Probability is the most important concept in modern science, especially as nobody has the slightest notion of what it means.

The participant-researcher held uncertainty, as measured by probability, as a central concept but retained variability as the central idea around which other ideas are explained.

Statistics is about asking and answering questions

Speed (1986) building on the work of Tukey (1977) presents yet another perspective: 'a major point, on which I cannot yet hope for universal agreement, is that our focus must be on questions, not models'. Data techniques and data, according to Speed, are to give us answers to more or less-well defined questions but statistics, it seems, has taken on another hue:

> …to summarise, display and otherwise analyse data, or to construct, fit, test and evaluate models for data, presumably in the belief that if this is done well, all (answerable) questions involving the data can then be answered. (p. 18).

Speed maintains that graduates in areas such as health and education working with government agencies, are called upon to answer questions, questions which arise in real world contexts. It is an understanding of the 'interplay between questions, answers and statistics' that is the basis of statistical thinking. Speed maintains that in order for students to understand problems in a real world context several issues must be canvassed. These included:

- drawing a distinction between statistical significance and significance to the discipline;
- how to deal with populations of 'real trains, field plots, cubic metres of ore' (p. 20);
- problems of selection and representativeness;
- identification of the population characteristics which are relevant; and,
- when should measures be standardised.

Speed's concern is that graduates are often not aware of multiple possible ways of using each given statistical technique (for example, regression) and how to marry uses with the research questions as they occur in real world context.

Probability, Design, Analysis and Interpretation

Cox (1997) presents statistics as a field of study that may be taken to have *three interlinked pillars*:

> 1. The mathematics of probability
> 2. General principles for design of investigations
> 3. General principles for analysis and interpretation.
>
> Formal principles of statistical inference are a part, but only a part of the third. (p. 261).

If a fourth pillar were to be added it would be, says Cox, 'so called official statistics' for government.

Concepts related to each of these perspectives can be identified in the maps provided by participants at Statistics '93. In a discussion of Cox's *pillars* or general principles of statistics, Hawkins (1977) refers to the 'differences of opinion within the statistics profession' about what constitutes *basic* statistical understanding and the extent to which this can be achieved without recourse to mathematical explanations. The participant-researcher's approach to resolution of this issue, of these differences in opinion, has been to adopt the approach that the most fundamental statistical concept should be one that provides coherence for the discipline. As a teacher (and practitioner), she is to seek systematically explanations for what she teaches in terms of this single concept, before drawing on additional ideas.

3.4.7 Other Experts: Conclusions

The analysis of the professor-expert's and other experts' statistical knowledge did not yield the definitive structure of statistical knowledge or even the definitive structure of the professor-expert's knowledge. As the search for concepts and the structure of statistical knowledge proceeded, the participant-researcher experienced a shift in perspective. The shift was from a cognitive approach to understanding knowledge (identifying knowledge, the concepts and how they were organised) to a constructivist approach (recognising many perspectives as to the nature of statistical knowledge) and finally to a hermeneutic approach (seeking the meaning behind terms used),

3.5 Unpacking the participant-researcher's expertise

What was evident in reviewing the professor-expert's and the other experts' concepts was a sense of similarity. The professor-expert's text could be described as Hawkins et al., (1992) described many texts: as following a *descriptive statistics-probability-inference* sequence. Schield (1996) suggested that introductory courses, of this ilk, are too narrow in focus, lacking real world breadth:

> When we think of statistics, most of us think of the subject as having several parts: descriptive, probability, inference and possibly modelling. But the essence of statistics is not readily determined by knowing these parts... The distinction is more than just a difference between theory and application. It is the difference in method (deductive versus inductive) and a difference in subject matter (formal probability theory vs. the material aspects of real data found in observational studies). (pp. 3-4).

The criticism of this sequence in teaching, and recognising the similarity in approach *data analysis-probability-inference* utilised by the professor-expert, created discord for the participant-researcher who had found the professor-expert's discussions so very illuminating. The participant-researcher again reviewed the work of the professor-expert in conjunction with the work of Hawkins. The differences between the professor-expert and others were attributed to the nature of the linkages between concepts, as illustrated by the discussion on measurement (refer Section 3.3.1), and to the ability to convey meaning not included in the definitions.

Therefore, in this section, the focus is upon:

- describing how concepts may be seen to take on different meanings (or different constructions of knowledge);
- examining the implications of adopting different approaches to probability; and
- examining the impact of data structures on the participant-researcher's statistical understandings.

3.5.1 A search for meaning

Between the two expositions, the case study and others, was the unpacking of the participant-researcher. At times the professor-expert would remark, 'you place too much emphasis on that idea or observation; it is not that important'. At other times, particularly in the classroom, distinctions were made that the professor-expert would not make. In this manner the unpacking and ownership of her own statistical expertise began. A search for a structure that could encompass the participant-researcher's statistical knowledge had begun.

At the outset, statistical theory was perceived as conveying *one true meaning*. In instances where the participant-researcher did as other statisticians did, there was now a questioning as to the interpretation of that doing, the meaning that was ascribed to the doing and the meaning that others ascribe to their actions. The issue of meaning was not about misperceptions, for which there is a great deal of literature (Falk, 1981; Hawkins, et al, 1992; Konold, Pollatsek, Well, Hendrickson-Amherst, & Lipson, 1990; Mevarech, 1983). The issue focussed on legitimate differences in meaning, incomplete meanings or multiple meanings. For example, Goodchild (1988) identified three types of meanings held by children for the arithmetic mean, those of location, expected value and representative number. Emphasis on one or other idea held has the potential to create a different meaning; similarly, choice of topics in subjects can be influential:

> In the more traditional areas of inference and decision there are lots of directions but they are all different... [Such difficulties] effect not only the choice of topics...[but also] the way we think about statistics. (Zideck, 1988; cited Hawkins et al, p.1).

Increasingly, as the research proceeded, it appeared that differences in meaning and interpretation may exist in relation to bodies of theory. As the participant-researcher found, new insight and new understanding of what was thought previously understood there was a questioning, 'What is it that creates meaning in the words, definitions and methods we provide for students?' For example, the participant-researcher did not always view the concept of variability as *the* fundamental concept in statistics. As suggested by their listing of fundamental concepts in the students' concept maps, these recent graduates also did not recognise this. Statistics *is* the study of variability. In adopting variability as the fundamental conception, suddenly the participant-researcher could create statistical theories and make meaning without having to rely on anybody else's definitions and rules. Prior to this, the participant-researcher felt no ownership of the rules, definitions and theory which were applied and interpreted. When the procedure was forgotten, there was no possibility of creating it: it needed to be retrieved from a manual or text. Meaning and ownership was experienced when the participant-researcher possessed a logic that allowed the reconstruction or creation of the theory and rules used. As Bliecher (1980) said:

> …meaning-full forms have to be regarded as autonomous, and have to be understood in accordance with their own logic of development, their intended connections, and in their necessity, coherence and conclusiveness; (p. 58.)

The issue of personal meaning, as distinct from statement of fact, arose in many contexts in relation to statistical concepts. Whilst working with the professor-expert and auditing a colleague's (statistical expert #3) class the participant-researcher found herself quietly asking, 'I know what I mean by that term: what do you mean by that term?' In finding meaning, the issue was not the *sense* of the words but rather their *significance* (Bleicher, 1980)*;* not the definitional nature of the term or combinations thereof, but rather in knowing one definition experts appear to call into play the implications of knowing that term.

The theory, philosophy and interpretation of meaning are the central concerns of hermeneutics (Bleicher, 1980). In establishing mutual understanding, Bleicher says:

> …people…establish mutual understanding…through the reciprocal mobilization of corresponding elements in the chain of their conceptual universe. (p. 54).

The ‘hermeneutical circle’ defines the problem, 'that whole in relation to which individual parts acquire their meaning’ (p. 2). For example, in the context of the comparison of concept maps of the professor-expert and those of the other experts, the listing of concepts may be similar, but the organisation of or emphasis placed on, one or other has the potential to create quite a different meaning. Discussions with the professor-expert often revealed such subtle difference in the organising component. It may have been a shift in the ordering of concepts or a difference in distinctions made. For example, when the participant-researcher spoke of three sources of variability, measurement error, sampling variation and chance the professor-expert used two, classifying sampling variation is simply an instance of chance.

The exploration of meaning began with probability, or more precisely, what it might mean to undertake statistics without probability. The exploration of meaning extended to examining how the participant-researcher identified analysis techniques utilising a *data structure* concept.

3.5.2 Why teach probability?

> I am sorry I cannot always use terms which are customary in statistics. A philosophical essay about statistical science ought to use the words already current. But we cannot safely use 'probability' for 'chance' because the word has been employed in too many different ways.(Hacking, 1965, p. 15).

The meaning of probabilistic statements came under scrutiny for the participant-researcher in a non-professional setting, when her daughter underwent a surgical procedure. The chances of one in a thousand, espoused by the surgeon, that something could happen, held little meaning. The sense, as a parent, was that of panic. When the risk was defined as akin to having a tonsillectomy that panic subsided, as this process had previously been successfully undertaken with another child. When that procedure failed to rectify the problem, and the next operation was quoted as having a 1 in 100 chance of some mishap, the parental reaction was that of total disorientation. What meaning can a probabilistic statement have for one, perhaps just a notion of more or less risk. A sensible understanding of risk for this child was impossible to grasp by the parent. The professor-expert suggested that perhaps this is a case of 'emotion over-riding knowledge.'

Coinciding with this sequence of events was the unpacking of the professor-expert, focussing on the topic *probability*. The joint paper (Micheal et al., 1993) and unpacking work with the professor-expert, focussed on visual representations of the theory in order to improve clarity and enable better understanding. What became an issue for the participant-researcher was the meaning or lack of meaning, which statistics held without recourse to probabilistic thinking. The purpose of addressing the issue 'Why teach probability?' was to identify the central position that probability plays in providing meaning to statistics.

Statistics and probability are often seen and taught as two distinct bodies of theory. Others view the two as enmeshed and describe them by the term *Stochastics*. Statisticians may take the need to understand probabilistic ideas as an imperative in order to appreciate statistical ideas or findings. However this same imperative is not perceived as such for all users of statistics; nor perhaps for the vast majority of teachers of statistics who were not educated from within the applied statistics or mathematics disciplines. Teachers of statistics come from many disciplinary backgrounds apart from statistics; backgrounds such as economics, education, sociology or psychology. As a consequence of these diverse backgrounds different orientations can be discerned in the teaching and the practice of statistics, orientations of which we must be aware and constantly critical.

Reflecting upon her early years as a social scientist, the participant-researcher asked 'how is it that more than a rudimentary level of probability knowledge is necessary if I, with my poor understanding of probability survived in the world of statistics?' Pursuing questions, so as to elaborate the nature of the relationship between probability and statistics, led to an early representation (refer Figure 3.37), and to an extensive pondering as to the nature of the links between probability and statistics.

Figure 3.37 The relationship between probability and statistics

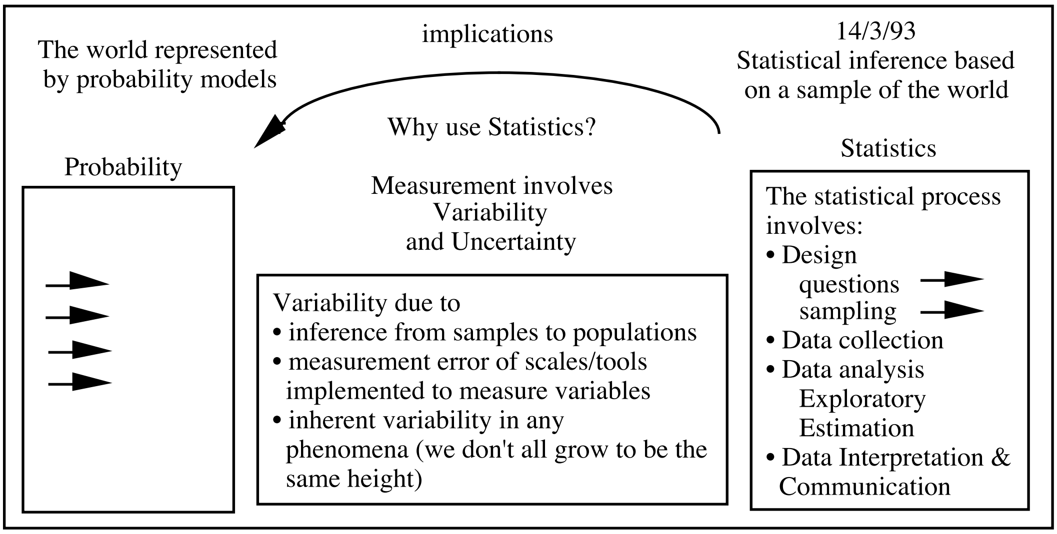


Figure 3.37 as drawn, led to the question: 'If one can understand the selection of statistical tests, manipulate them appropriately and recognise when there is a significant difference, then does one need to expand the probability side of this representation at all?' (speaking as a user of statistics, not as a statistician). How is it that one can use statistics from a non-probabilistic framework?

Reflecting upon the participant-researcher's early days as a social scientist, statistical questions then asked came to the fore, as did the interpretations drawn from the attendant analyses. The textbook by Seigel and Castellan (1988) can be used to illustrate how the participant-researcher functioned statistically. What was distilled from this invaluable text was a sense of data structure, as described in Table 2.1, and a recipe for testing a given hypothesis (refer also, Johnson, 1981) as follows:

Step 1: Specify the null hypothesis and alternate hypothesis

Step 2: Determine the appropriate statistical test given the data structure and measures

Step 3: Decide upon the alpha level

Step 4: Determine the appropriate sampling distribution

Step 5: Determine the rejection region

Step 6: Make a decision

A typical question might have been, 'Is there a significant difference in the mean cholesterol levels after treatment with the drug known generically as gemfibrozil?' The nature of the interpretation may be as follows.

Interpretation 1

Using a matched pairs t-test a significant difference (t=5.489, df=191, p=0.0001) in mean cholesterol levels was found between measures at the commencement (mean=6.24; sd=1.22) of treatment with gemfibrozil and after the treatment regime had been implemented for three months (mean=5.89; sd=1.06).

This approach effectively absolves the researcher from making a decision as to the meaning of the results; the results are significant. If $p < 0.05$ then a significant difference is found. What is concealed by this approach is the fact that statistics is not deterministic. This first interpretation led to the following.

Interpretation 2

The probability of obtaining a difference of this magnitude (mean difference=0.35) in the mean cholesterol (with the pre-treatment (mean=6.24; sd=1.22) and post treatment (mean=5.89; sd=1.06) is 0.0001 given a true null hypothesis that the difference in means between the two

> treatment groups is zero. This means that it is not very likely that the difference would occur in chance samples in which the members had not been chosen in a special way or given special treatment. It can be concluded, with some confidence, that the treatment with gemfibrozil was effective.

What is not made explicit in the first interpretation is that the statistic does not provide *the answer,* it provides only a part of the answer, for:

> [a]s long as students believe there is some way that they can 'know for sure' whether a hypothesis is correct, the better part of statistical logic and all of probability theory will evade them. (Konold, Pollatsek, Well, Hendrickson-Amherst, & Lipson, 1990, p. 92).

One needs to understand the probabilistic nature of statistics and as Hawkins et al (1992), suggest, a 'curriculum and a pedagogy which break with traditional (deterministic) mathematical values' is necessary. Statistics, as dominated by the classical approach wherein fixed levels of significance (alpha=0.05 or 0.01) are used to determine whether or not an hypothesis is accepted or rejected, is unwittingly contributing to this deterministic usage.[3] What meaningful difference can there be in obtaining a p value significant at an alpha of 0.05 as compared to alphas of 0.049 or 0.051? Of course the situations statisticians have to deal with become a little more complicated. Consider the situation where a multivariate test is not found to be significant at 0.05 but various univariate tests are found to be significant. Does one stop at that magical point where $p>0.05$ or does one continue to explore the univariate results? Here, differences can be found between statisticians.

The questions posed on the statistical side of Figure 3.37 often do not make their probabilistic basis explicit, as discussion of the gemfibrozil treatment interpretations suggested. The generalised form of questions as posed on the statistics side:

- are there differences between groups? and,

---

[3] Indeed the setting of the alpha level at 0.05 or 0.01 is a reflection of the limitations of 'computing power' in the 1920's. An artifact of the times, fixed alpha levels levels have become accepted scientific convention and this convention is still more readily observed in scientific publications than the more informative alternative of providing the p-values.

- are there any relationships between variables?

These would be appropriately rephrased as:

- What is the probability of attaining a difference of this kind between the groups if there is no real difference? and,
- What is the probability of obtaining a correlation coefficient of this size if there were, indeed, no relationship?

These, of course, were questions of the hypothesis testing kind and do not address estimation. As the professor-expert suggested too often there is a tendency to test hypotheses rather than to estimate parameters. The earlier question and interpretation, regarding the effectiveness of the gemfibrozil, can be addressed by estimating the difference in cholesterol levels before and after as follows:

Interpretation 3

> The 95% confidence interval provides an estimate of the difference in the two means, before and after treatment and attaches to this an interval. If this procedure were conducted 100 times then we would expect 95% of the confidence intervals to cover the true mean difference between the two treatments. In this particular instance the 95% confidence interval of the differences in means (mean difference=0.345, df=191, sd=0.871) is 0.221 to 0.468 mmHg. This is just one of the possible confidence intervals that could have been calculated if the experiment were repeated.

The questions and answers obtained in the application of inferential statistics are probabilistic and this implies that the concept of probability should be understood in order to understand statistical results. Unfortunately, many interpret the outcomes of scientific method deterministically - there is or is not a significant difference. What is lost is the statistician's imperative to understand the idea of chance. In their statistics text, Moore and McCabe (1993) assert that the most fundamental probabilistic idea is that:

> Chance is all around us. Sometimes chance results from human design, as in the casino's games of chance and the statistician's random samples. Sometimes nature uses chance, as in choosing the sex of a child. Sometimes the reasons for chance behaviour are mysterious… (p. 279).

Pfannkuch and Brown (1996) provide an example that can be used to illustrate chance happenings:

> Traffic Accident Question: On average there are 600 deaths due to traffic accidents each year in New Zealand. A person observed the following:

		Number of Deaths
February	Week 1	3
	Week 2	12
	Week 3	21
	Week 4	14
March	Week 5	2

> Assume that none of these weeks contain a holiday weekend. Suppose the headlines in the newspaper claimed that week three was a 'black' week and police reported that speed was a factor. The next week was described in the papers as more evidence that New Zealand driving was deteriorating. At the end of week five the police congratulated themselves for the low death rate -- their extra patrols had succeeded. What would you say to this person? (online).

A similar example was provided in the Australian media in the early 1990s and this formed the basis of a discussion between the participant-researcher and the professor-expert. The report was to the effect that the police were alarmed that the road deaths over a particular weekend had increased from, say, 5 to 13 deaths on the same Easter weekend in consecutive years. This increase was attributed to drivers reverting to the old ways. The professor-expert provided an answer to the question, 'to what can the change be attributed?' Paraphrasing his response, the answer provided was that this question requires an understanding of probability; this notion of chance. The change may be due to many factors including chance. These two figures may simply represent two independent realisations of some random process. To have any conception as to the unusualness or not of the change one possibly should also have additional information regarding the nature of the data. Such additional information might include deaths in the previous years, for example, the knowledge that the five road deaths last year was the lowest in the preceding 20 years or the incidence of multiple fatalities in a single accident. One also needs some knowledge of probabilistically based models, including the probabilities that a Poisson random variable with a mean of say 9, takes on the values 5 and 13 or even more extreme values. In this instance, the probability in each tail exceeds 0.1.

In the instances described, a change in the understanding of the relationship between probability and statistics (the study of variation), led to an appreciable change in the interpretation, and hence meaning of the statistics calculated. However, this is not to suggest that meaning is congruent with understanding. Hawkins et al (1992) concur with Margenau (1950) as cited by Van Brakel (1976):

> General discussions of the meaning of probability by philosophers have lately shown little evidence of agreement upon any common view, and the literature is becoming progressively more confused. (Hawkins et al, p. 69).

One of the problems arises where probabilities:

> … are based on a combination of subjective beliefs and relative frequencies. The mathematical method of combining the two uses Bayes' theorem, but people do not always adjust probabilities according to Bayes' rule. (Hawkins et al, 1992, p. 69).

According to Hawkins et al, the solution for some teachers teaching probability has been to adopt a mathematical approach:

> … beginning with a set of axioms (common to any interpretation of probability) and formally constructing deductions from those axioms...there is a price to pay for taking this route - namely, that results based on mathematical probability may clash with intuitive notions... (p. 70).

Perhaps it is the case that the unenlightened do not understand, but the area of probability is one which, to this day, is in the throes of debate where different interpretations or subjectivity exist. Further development of the probability topic is beyond the scope of this thesis. Suffice to say that without a real understanding of the concept of probability it is impossible to come to grips with the statistical world in which we live – one of variability in the face of uncertainty.

3.5.3 Data Structures

The term *data structure* is somewhat nebulous and, as used by the participant-researcher, is not to be confused with *structure in data*. Bishop, Fienberg, & Holland (1975) use the term *data structure* in the second sense when they describe investigators attempting to explore data of unknown structure by fitting parameters to a model in order to attain a good, but not saturated, fit. They also exemplify how structure is imposed on data, the concept of data structure, and that this is linked to analysis:

> When the cells are defined in terms of the categories of two or more variables, a structure relating to the nature of the data is imposed. The natural structure for two variables is often a rectangular array with columns corresponding to the categories of one variable and rows to the categories of the second variable; three variables create layers of two-way tables, and so on ... A good mathematical model should reflect this structure. (p. 10).

When one analyses data it may also be to find structure; such as location, spread, distribution, clusters, unusual data or dependencies within the data. Data structure as used by the participant-researcher, is more akin to a description of design features, although there is some overlap (for example dependencies are also a feature of design). As used by the participant-researcher, data structure is a concept to do with the recognisable structure of data before analysis. It is what helps define the analysis that should take place and it may be modified as the process of analysis begins. Structure in data is structure found through analysis.

The professor-expert used the term *data structures* - 'there are many types of data structures' (Griffiths et al., 1998, p. 140) and these include:

> …the simplest data structure contains a single measurement (observation) on a single characteristic (variable) on each of several 'items' (individuals, plants, towns, components, days, ...) A data set of this type is called a univariate (or one variable) data set. (p. 14).

The concept was expanded to include multivariate data (or many variables) and time series where measurements are made at successive intervals of time. Questions emerged when unpacking the professor-expert.

- What terms did the professor-expert use to distinguish data structure?

- Did the expert use the type of measurement (nominal, ordinal, interval and ratio, or discrete and continuous)?
- Was the concept of measurement scale used in in addition to the number of characteristics measured (univariate, bivariate and multivariate)?
- Was the nature and number of batches of data (related or independent; one, two or more) used to distinguish between different data structures?
- Were there other features of data that the professor-expert included in the data structure concept?
- Were there other possibly more complicated data structures?

The term data structure was not indexed in the text. It does not appear to have a formal definition. However, the concept of data structure, because of the described linkages of specific exploratory data techniques and statistical techniques to specific data structures, was seemingly of some importance and this was confirmed by the professor-expert.

Other users of statistics also convey 'the sense of data structure' as being important (Friedman, 1972; Hindle, 1993; Howell, 1995; Seigel & Castellan, 1988; Twaite & Monroe, 1979). Each of these has a table or figure aiding the reader in the selection of statistical techniques. As with the professor-expert, the concepts of univariate and multivariate are important. Many classifications of data, for example, nominal, ordinal and interval (Seigel & Castellan, 1988; see also the manual for the statistical software JMP[4]), the classification qualitative, quantitative, continuous and ranks (Howell, 1995) and nominal, ordinal and continuous (Hindle, 1993). This latter classification is also used in the actual statistical software for JMP. It is a classification that fails to distinguish between discrete and continuous data. Other concepts include related and independent variables (Seigel & Castellan, 1988), matched and unmatched data (Hindle, 1993) and dependent and independent variables (Howell, 1995). The nature of the question asked (relationships or differences) and the number of independent variables are also used by Howell.

---

[4] Registered Trademark, Product of the SAS Institute Inc. (SAS Institute,1989)

Many statistical packages such as GENSTAT (Payne et al., 1987), JMP[5], SPSS (Nie, Hull, Jenkins, Steinbrenner, & Bent, 1975) StatView II (Feldman, Hofman, Gagnon, & Simpson, 1987) and SAS[6] also use concepts regarding data structures. In these instances, they are a mix of computing (for example, reflecting data input and retrieval or data type classified, for example, as numeric, text, and date) and statistical design concepts. Three examples have been chosen to illustrate how statistical software packages appear to be using data structures concepts in a statistical sense. These three are JMP, SPSS and GENSTAT.

JMP

JMP provides a tabular summary of statistical techniques to aid in the selection of a technique for analysis. Choices are made by distinguishing between roles (response and explanatory variables and the number of each type, of one or more), which each have their own measurement (nominal, ordinal and continuous). It uses the concept labelled platform, which, in a limited fashion, addresses Speed's *different questions*, for making choices between different analyses, for example, description, testing means or fitting models.

SPSS

SPSS provides a general form for specifying each statistical procedure. The structure provided by these procedures mirror statistical design features. For example, the general form of a single analysis of variance procedure is as follows:

```
ANOVA   dependent variable list BY independent variable list
  (min,max)...[WITH covariate list]
```

This has embedded within it data structure concepts such as variable type (dependent, independent and covariate), data measurement (through the minimum, maximum codes relating to nominal categories and continuous, or perhaps, interval/ratio data embedded

---

[5] Registered trademark

within the definition of covariate). The covariate list is optional allowing one to address different questions with the analysis, that is, the effect of the independent variables (predictor variables) upon the dependent variable (response variables) without covariates considered, or after variation due to the covariates has been removed. More complicated designs such as multivariate analysis of variance (MANOVA) use additional structure. The responses (multiple dependent variables) are expected to be statistically dependent. Other structural issues, such as what to do when data are missing, are also incorporated into these procedures.

GENSTAT

GENSTAT defines data structure, in the computing sense, as the structure within which data are stored. Some of the computing data structures include pointers, matrices, text, decimals and vectors. Text may be used in a fashion similar to nominal measurement. Other structures are akin to statistical data structures, for example, variates 'contains a list of numbers' (p. 59) or factors which 'indicate groupings of units' (p. 67). Computerised statistical procedures involve the retrieval and use of data. To do this, the data structures must be mapped in some way to different statistical designs. Whilst examining GENSTAT, the simple statistical notion of data structure, used by the participant-researcher, merged with statements about complex design features. GENSTAT deals with designs that are more complicated than those permissible in most other packages. More complicated design to the participant-researcher is akin to more complicated data structure. Terms, such as blocks and treatment structures, balanced designs, nested designs, incomplete designs describe designs of varying complexity (that is, data structures).

In a statistical sense, the term *data structure* is not endowed with a formal definition in the standard literature. Perhaps it should be left as the coming together of many concepts, yet it was the holistic notion of data structure that the participant-researcher considered essential to her functioning as a statistician (the doing component). Given

---

[6] Registered trademark

the professional and educational development path that the participant-researcher has followed - educational psychology, computing, statistical-computing to statistician - it is not unusual that the data structure concepts from different disciplines have merged and perhaps mutated. One of the purposes of this thesis was to make explicit those implicit understandings. An example drawn from thesis work (Porter, 1987), supervised by the professor-expert, provides an illustration of several aspects of the data structure concept.

This thesis, *Models for Ordered Categorical Data,* examined common forms of analysis of a particular data structure, such as presented in Table 3.4. The study involved researching drug usage in two cohorts of students. In each cohort there were two groups, a control group and experimental group that received a drug education program. Two features of the data are important when determining what form of analysis to conduct. The *response* variable was measured on an *ordinal scale*. The *one or more stimuli* (treatment versus control and year) were measured on a *nominal scale*. The question of interest would be whether or not there was a *co-occurrence* between the response and two stimuli. The data structure is deceptive. Data of this form tend to be incorrectly analysed, using either a chi-squared analysis or an analysis of variance (ANOVA).

Table 3.4 Ordered categorical response data: Proportion of responses to the question: 'On how many days in the last four weeks did you have an alcoholic drink?

Student	Group	None	1-2	3-5	6+	Total N
Year 7	E	.76	.19	.05	.01	217
	C	.42	.44	.13	.02	115
Year 9	E	.47	.33	.15	.05	129
	C	.42	.38	.14	.06	111

E=experimental C=control

Why is chi-square inappropriate? As Tukey (1954) asked:

> How soon will we appreciate that the columns or rows of a contingency table usually have an order? When there is an order, shouldn't we take this into account in our analyses? How can analyses be efficient otherwise? (p. 724).

The application of the chi-squared test does not use the additional information regarding order in the data: it treats the response variable as measured on a nominal scale. For the appropriate use of a chi-squared test it does not matter whether or not the columns are interchanged in order. To conduct an ANOVA involves assigning values 'none' (0), '1-2' (1), '3-5' (2), '6+' (3) for the response variable which was actually measured on an ordered scale. A further complication is that these data are actually based on an interval scale which has been collapsed/summarised on to an ordinal scale. That is they are grouped quantitative data, a 'stronger' structure than ordinal data which is, in turn, stronger than nominal data. It is, at least in principle (and in practice if it is important enough) possible to model the explicit structure of grouped quantitative data. Grouped discrete data are different to grouped continuous data. The subsequent treatment of these data as interval data, (perhaps implicitly with a normal distribution), involves attaching different structure to the data than is actually present. The use of ordered categorical analysis involves the examination and modelling of the cumulative probabilities, thus overcoming the problems associated with both inappropriate modes of analysis (chi-square and ANOVA).

Other data sets, which look similar, may not be suitable for analysis of this kind. These deceptively similar, but distinctly different, data structures, involve one or more of the following situations, as illustrated in Table 3.4.

- the stimuli are measured on an ordinal scale (drinks: 'none', '1-2', '3-5', '6+') rather than on a nominal scale.
- the 'raw' response data are ungrouped, quantitative data. In this case it preferable to perform an analysis which utilises all the information in the data, rather than to group the data.
- the data are presented as percentages (and the row the count is unavailable), and are thus not multinomially distributed.
- the observations in different rows are not independent, as would happen if a longitudinal study (same people) were used in Year 7 and 9, rather than a cross-sectional study at a particular time.

In making explicit the understandings that have served to enable the participant-researcher to function effectively as a statistician, exploring the ever-developing concept of data structure has been essential. The participant-researcher was not a specialist in design (nor are many statisticians) and yet design appears to hold the key to a fully elaborated notion of data structure. How can an incomplete notion of data structure be useful?

There are at least two ways this incomplete knowledge of structure may be useful.

1. It can be used to identify data structures which fall within it (techniques suggested at which point more detail can be gained) and those which appear to fall outside it (and for which a search of the literature will be warranted to see if the structure fits what has been described elsewhere, or some real statistical work to develop appropriate techniques). It is similar to the process Bishop et al., (1975) describe:

> By juxtaposing practical examples from a variety of fields, the researcher can gain insight into his own problem by recognising similarities to and differences from problems that arise (p. 2).

2. The most commonly used analytic techniques tend to fall within the partial structures whilst more complicated designs fall outside the framework. At the design phase of research this might provide a useful indicator that a better design (for example, balanced rather than unbalanced) might be warranted.

To see the data structure, to make meaning from it involves knowing more than the definitions of types of data (nominal, ordinal, interval/ratio), or type of variable (response, explanatory or stimulus), or numbers of variables analysed (univariate, bivariate and multivariate) or what characterises the dependencies between variables. To see the data structure, and hence *target* appropriate forms of analysis, all the concepts related to data structure (including those characterising more complicated design, such as balance, nesting, blocking) must be activated and viewed in conjunction with each other.

The expression *target the appropriate analysis* is deliberate: the intent is to focus on a range of possible options rather than deciding on a single analysis technique. The use of measurement information provides a starting point in the selection of a tool for analysis. In the example provided in Table 3.4 the data are drawn from a discrete distribution and since it is the number of days out of 28, the binomial is the natural first choice of model. However, since people (including young people) drink more on some days of the week than on others, the binomial is probably too simple. Some discrete distribution with an upper bound will be appropriate. It may be that the Poisson is a fair 'model' with an almost zero probability beyond 28 days. Only through the process of fitting the models will inadequacies of these assumptions be highlighted (for example, cell residuals might be too large indicating a lack of fit for that part of the data). In this manner probabilistic concepts also come into play. What is fortunate for statisticians, is that various forms of measurement are often reflected in common probability distributions such as the normal distribution. However, it is also the case that a number of distributions may provide a reasonable fit to the same set of data, hence requiring selection of a model for other than statistical reasons. When an appropriate probability distribution is used, knowledge regarding the rarity or otherwise of particular statistics is available and this is useful for hypothesis testing.

Intertwined with the data structure is the nature of the question asked, for it is the combination of the two that determines the analysis technique and the precise manner in which the technique is implemented. Speed draws attention to the earlier work of Mosteller and Tukey (1977) reminding that most techniques, including regression and times series, have multiple functions. Different implementations of statistical procedures address different questions. The aims for regression analysis are summarised as:

1. to get a summary;
2. to set aside the effect of a variable;
3. as a contribution to an attempt at causal analysis;
4. to measure the size of effect;
5. to try to discover a mathematical or empirical law;
6. for prediction. (Speed, 1986, p.24).

To develop further the data structure concept involves further elaboration of design features. However, this is beyond the scope of an introductory course and, as such, is beyond the scope of this thesis.

The sense of data structure and a process for the conduct of research have provided the participant-researcher with the where, when and how to conduct quantitative research and to analyse it. Probability theory has imbued decision-making and interpretation with a probabilistic rather than deterministic flavour. It is, however, an understanding of the nature and cause of variation, together with measurement concepts, that has provided the participant-researcher with an ownership of ideas and the capacity to build or challenge statistical theory. Hence, the theme 'statistics is about variability' became the focus of the statistical literacy curriculum implemented in the second phase of this research.

3.6 A postscript: Where is variability?

> Statistics is a methodological discipline. It exists not for itself but rather to offer to other fields of study a coherent set of ideas and tools for dealing with data. The need for such as discipline arises from *the omnipresence of variability*. Individuals vary. Repeated measurements on the same individual vary. In some circumstances, we want to find unusual individuals in an overwhelming mass of data. In others, the focus is on the variation of measurements. In yet others, we want to detect systematic effects against the background noise of individual variation. Statistics provides means for dealing with data that take into account the omnipresence of variability.(Cobb & Moore, 1997, p. 801).

This very clear, explicit statement, one might suggest the fundamental principle underlying statistics, is used to discuss how statistics should be taught. By the concluding phase of the research the participant-researcher had begun to suspect that, even if variability were explicitly recognised as the central concept by statisticians, that *the language of the discipline would obscure the central concept, variability. Furthermore, this obscuring of the concept of variability is particularly pronounced in the textbooks which are written for students*. For example, Cobb and Moore, made the explicit statement to their professional fellows that the need for 'the discipline arises from *the omnipresence of variability*'. However, neither of the indices of two introductory texts co-authored by Moore, *Statistics, Concepts and Controversies*

(Moore, 1991) and *Introduction to the Practice of Statistics* (Moore & McCabe, 1993) has an entry for *variability* or *variation*.

A ministudy of index references to variability

A ministudy was used to examine the extent to which textbooks more generally make an explicit index reference to *variability* or *variation*. Most textbooks provide an index. McKechnie (1957) provides three definitions for an index:

> 1) that which points out, a title, from *indicare*, to point out, indicate,
>
> 2) (a) an alphabetical list of names, subjects, etc together with page numbers, usually placed at the end of a book or publication; (b) a list describing the items of a collection and where they may be found; catalogue, or
>
> 3) a table of contents, preface, prologue, or statement of subject. (p.929).

One might expect that the fundamental statistical concept, *variability* or *variation* would be indexed or that the table of contents would reveal a chapter heading devoted to the topic.

The sample

To examine the proposition that textbooks do not make explicit use of variability or variation as an index reference, the name of textbooks electronically listed under the subject description *Statistics* were drawn from the University of Wollongong Library catalogue. The textbooks that appeared, from the title, to represent advanced statistical subject matter, or not relating to the teaching of Introductory Statistics were deleted. Texts written for the statistics disciplines from the perspectives of Commerce (Accounting, Economics, and Management), Behavioural Sciences (Psychology, Nursing, and Education) or Statistics (no discipline reference) were included: other texts were excluded. A comparison of the explicit representation of variability in the index amongst the three broad discipline areas was anticipated, but numbers too small to be useful. A comparison was therefore made between the Statistics discipline and the others combined. The texts were classified according to date of most recent publication. For any book with more than one edition, only the latest edition was included. Samples were drawn from the decades, 1970s (n=41, 61.2% of books listed), 1980s (n=35, 62.5% of books listed) and 1990s (n=20, 62.5% of books listed). The samples reflect the number of items originally listed; these declined from the 1970s to the 1990s. Only books available on the shelves, and found over the two days that data were collected, were examined. This may have biased the study against 'popular' books, more likely to be on loan or moved to the closed reserve section of the library. The sample itself has the further inherent bias that derives from the collection itself being purchased over the years following requests from a small number of academics, from several disciplines, with an interest in statistics.

Measurement

The index of each text was examined. Whilst the initial classification was 'no entry' versus 'entry', as the references were examined similarities and differences in the nature of the entries began to emerge. Where variability or variation was indexed the nature of each reference was examined to ascertain a more complete description of the entry than revealed by the index alone. Three classifications were eventually used:

- no entry;
- measures of variability, which generally included a definition of variability in the measurement context (and, in one instance, simply the coefficient of variation and another variability measure defined in the context of ANOVA); and
- multiple representations of variability drawing from definition, measures, graphical displays, questions in relation to variability, sampling, sources of variation (chance, biological, inherent), variability associated with ANOVA (within, between, total, explained, unexplained), variability associated with time series (seasonal, cyclical, temporal, systematic), variability in relation to design, and variation in relation to quality control.

Results

Of the 108 texts that were examined, 39.8 percent (n=43) did not index the terms variability or variation, whilst only 21.3 percent (n=23) provided multiple representations of variability. Books that were classified as general statistical, rather than commerce or behavioural sciences, were no better in this regard with 45 percent of these revealing 'no entry'.

The breakdown for the decades 1970s, 1980s, and 1990s of the proportion of texts that had an explicit reference to variability or variation is provided in Table 3.5. The entries are reasonably consistent from decade to decade. Small decreases in the proportion in the 'not entry' category and similar small increases in the 'multiple entry' category are not statistically significant.

Table 3.5 Index entries for variation or variability

	1970s n (%)	1980s n (%)	1990s n %	1970-1990s n %	Statistics only 1970-1990s n %
No Entry	18 (42.9)	14 (40)	11 (35.5)	43 (39.8)	34 (45.3)
Define & Measure	17 (40.5)	13 (37.1)	12 (38.7)	42 (38.9)	25 (33.3)
Multiple concepts	7 (16.7)	8 (22.9)	8 (25.8)	23 (21.3)	16 (21.3)
Totals	42	35	31	108	75

Partway through the classification of entries it became evident that there were subtle but potentially important differences within the categories *define and measure* and *multiple concepts*. Whilst there was no attempt to count the different instances for the sample that had already been classified, it was thought prudent to characterise some of the references to variability, for it is these references that students will be subject to in reading their texts.

The term *variability* was often introduced to students in the context of *measures of variability*. However, in this context differences were observed in the manner in which variability was *defined* and *exemplified*. Some authors (Lapin, 1990; Roscoe, 1975) lead into measures of variability through an introduction to the inherent variability in natural phenomena:

> Alone, any value of central tendency inadequately summarizes the population. It does not indicate how different from each other the population values are. Practically all populations exhibit variability. Nowhere is this more true than nature. Just as persons come in different shapes and sizes, so do all natural phenomena vary in important characteristics. There are wet years and dry years. Earthquakes occur with varying intensity. People vary in their talents. Economics growth is unsteady and corporate profits fluctuate -often widely over time. (Lapin, 1990, p. 52).

However the most common way in which students are introduced to variability measures is through the concept of spread (or dispersion), without the context of variability being inherent in natural phenomena. This is variability was located in the data collected or simply in a set of *numbers without context.*

> The statistics described in this chapter indicate the spread of scores in a distribution. They are called measures of variability.(Jaeger, 1990, p. 3).

> One of the most important characteristics of almost any data set is that the values are not all alike…We require ways of measuring the extent to which data are dispersed, or spread out and the statistical measures which provide this information are called measures of variation. (Freund & Simon, 1992) p. 53.

> The term 'variation' means dispersion, scatter or spread. A measure of variation sums up in one figure the amount, or degree, of scatter of the values of the group so that we can determine how well an average represents the data from which it was computed. This measurement is usually made of the variations of the values from their arithmetic mean… (McElroy, 1971, p. 85).

Any book that only has an entry for 'measures of variability' is not really getting to grips with the broader importance of variability (eg for understanding variability, partitioning variability, displaying variability in data, designing experiments to control for unwanted sources of variability).

Another typical introduction, a variant of the last, leading to measures of variability is that of variability being another important feature in addition to the mean (Downie & Heath, 1974; Heitzman, 1980; Ott & Hildebrand, 1983b). For example:

> Averages locate the center of a distribution but tell us nothing about how the scores or measurements are arranged in relation to the center… (Downie & Heath, 1974, p. 49).

> The three measures of central tendency - the mean, median and mode - locate the center of a distribution of data, but they tell you nothing about the spread or variation of the

> measurements…Have you ever heard the story about the statistician who could not swim and drowned in a river with an average depth of 3 feet?…it does stress the importance of variation. (Ott, Larson, & Mendenhall, 1983a, p. 109).

This form of introduction is sensible only in that, for most data sets, there is an interest in describing the centre, the spread and the shape of the distribution. On other occasions, variability is introduced as a pattern called the distribution of a variable:

> Our interest in collecting data is to detect a pattern of variation in the values of a variable. Such a pattern is called the distribution of that variable. (Krishnamurty, Kasovia-Schmitt, & Ostroff, 1995, p. 112).

Some 38.9 percent (n=42) of texts reference only these basic ideas about variability. In the category of *mutiple entry* these ideas are often extended by a single additional theme, such as variation associated with sampling or design. A few of the multiple entries provided richer references to variation. These might include references to inherent variability in phenomena, variation in measurement, measures of variation, variability associated with design and in this lone extract a reference to variability as, *organized, following a prescribed pattern as an introduction to random variables*:

> We have noted in Section 2.3 that variability is present in our measurements…variability occurs between the height measurements made by two different individuals… Variation is universal in characteristics of all populations. We live in a variable world…Since variability is universal we must learn to live with it and to design experimental and survey investigations in such a way as to overcome its effects. Some types of variability occur in an organized fashion (Chapter 12) as a result it is possible to devise statistical procedures for summarizing the sample facts and… [leading to measures of variation] (p. 36).

> In certain situations, we can go further and define variability completely when we know that the observations follow a prescribed pattern, that is, the variability is organized in to a pattern rather than being completely unknown…many types of organized variation have been described mathematically. (Chapter 12, Federer, 1991, p. 394).

The entry by Glass & Stanley (1970) is used as an introduction to measures of variation, but draws on design to reduce and explain variability:

> Other descriptive measures are required to measure the variation of scores within a group. In this chapter several statistics that measure, in different ways the variability (heterogeneity, dispersion, scatter) in a group of scores will be presented and discussed. Later in this text, you should begin to see that some of the most important functions of statistics relate to procedures by which variability, which is in a sense uncertainty can be

> reduced, explained or accounted for. The whole scientific enterprise is concerned with notions of variability. When much unexplained variability exists, predictions can never be accurate. But when explanations can be fashioned for why people or things differ, uncertainty can be reduced and portions of variability can be removed. (p. 75.).

Kuzma (1992) is more explicit in the entries, tapping the measurement of variability, design to control variability, and measurement error:

> Knowing a distribution's central tendency is helpful, but its not enough. It is also important to know whether the observations tend to be quite similar (homogeneous) or whether they vary considerably (heterogeneous). To describe variability, measures of variation have been devised. (p. 48).

> Understanding the sources of variation may help you appreciate the meaning of standard deviation. For example, among subjects, one source of variation may be due to a personal characteristic such as age or sex. Another source of varying conditions of the subject (that is observations obtained before or after dinner, or before or after exercise, may differ). Yet another source of variation is measurement error. Although a certain amount is inherent in any observation, scientists strive mightily to keep it to a minimum by use of appropriate experimental designs. (p. 51).

Why are the concepts *variability* and/or *variation* comprehensively indexed in so few of these books? Why is variability not highlighted, boxed, along with other important concepts as often happens throughout texts for other disciplines or even for concepts such as the standard deviation in statistics? When statisticians look at data, they may see variability, but they talk in terms of patterns, outliers, trends, centre, spread, maxima, minima, gaps and clusters. When statisticians speak of the variability attributable to sampling, they speak in terms of sampling error. For example, Howell's sole index reference to variability is 'variability due to chance' (1992, p. 692). The actual reference says merely:

> Along the same lines, the phrase **sampling error** is often used in this context as a synonym for variability due to chance. (p. 80).

Similarly variation in measurement may be spoken in terms of measurement error, not in terms of variation.

Some statisticians refer to this phenomenon, of knowing variation by other descriptions. For example, Mosteller, Fienberg, & Rourke (1983) speaks about the term, *the differences between means*:

> In the ANOVA, we measure variability and allocate it among its sources. We have already seen such analyses in the tests of differences between two observed means. For that problem, the sources of variation were (a) the difference between the populations means and (b) the variation with the two samples. We used the variability within the samples to assess the size of the observed differences in means, that is, the variability between samples. We use these same ideas of variability within and variability between… (p. 436).

> As we proceed with ANOVA the reader may feel we are analyzing means instead of variances. But it is the variation in means that draws our attention. We want to know how large that variation is and whether it is important or can be neglected. (p. 440).

In his index Zuwaylif (1984) similarly alerts students to multiple terminologies:

> Explained variation. The amount of variation in the dependent variable Y that can be explained by changes in the independent variables. The explained variation is sometimes called variation due to regression or sum of squares for regression. (p. 600).

> …The unexplained variation is sometimes called the residual variation, variation due to error or residual sum of squares. (p. 607).

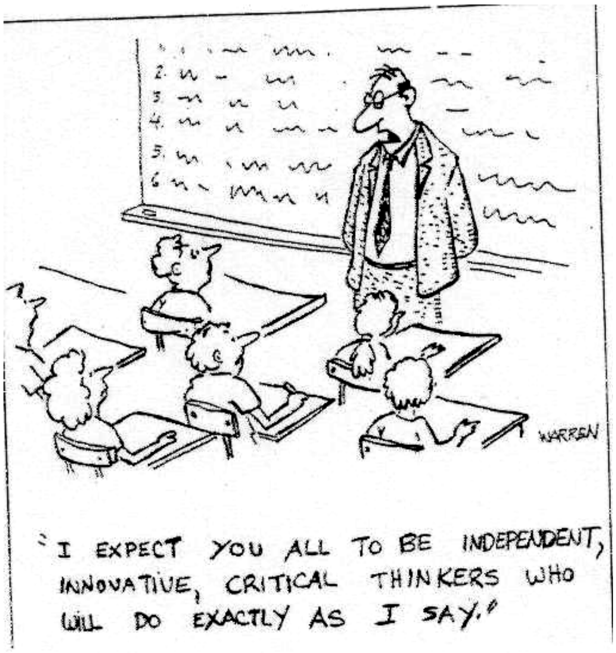
It is generally left to the student to connect the multiple terminologies and recognise that they are dealing with aspects of variability.

Conclusion

The examination of the text indices suggests that it is plausible that texts do not create an explicit focus on variation as the most important statistical concept. The textual material does reveal references to variation not detected by the index. However, one must wonder whether students attend to these references or whether they focus on the definitions and titles that are highlighted. It seems likely that this *omnipresent variability*, is so pervading in its presence, obvious in meaning from everyday language, that it is experienced primarily through a multiplicity of other terms. Schacht (1990) reviewed 12 statistics textbooks written for students in social and behavioural sciences and published in 1990. Variability was listed as a topic in these textbooks only in the context of being a *measure of variability*. A more extensive analysis of textbooks, contents, texts, maps and frames, is necessary to determine how variation is represented and how it can be better represented as an organising principle. This is beyond the scope of this thesis. However it provides a plausible next step in understanding the difficulty students confront when learning statistics. In the context of this thesis I would contend that it is not just experts and/or teachers that need to make their knowledge explicit. In the case of statistics, the discipline needs to make variability explicit through its texts, teachings and resources.

SECTION 3

Implementing curriculum



Phi Delta Kappan, Vol 71(1) 1989, p. 38.

CHAPTER 4

Preparing To Teach

4.1 Introduction

The *raison d'etre* for this work has been to develop a grounded theory of the teaching and learning of statistics. The process of unpacking and repacking statistical expertise, described in Section 2, was undertaken to make explicit the teacher's and other experts' statistical knowledge. This was to maximise the chance that, in the statistical classroom, the *what* and *why* of the teacher's statistical theory were not replaced simply by the procedural how. It was also used to determine the most fundamental statistical concepts, so that they could form the focal point of teaching.

In this second phase of research, the aim was to develop an exemplary course in statistics. This was to be achieved through the cyclical processes of teaching, reflecting, further unpacking of expertise and searching for explanations of what was observed in the classroom. What the description of an exemplary course might entail varies with circumstance. The teaching in this study was to take place at the tertiary preparation level. Thus, whilst there was a requirement for students to master a number of statistical concepts, there was also a need to position the student as academically able, perhaps enthusiastic but certainly a confident learner of statistics. As an educator and statistician, one could also suggest that, at any level of study, the ability to understand statistical concepts is more desirable that the blind application of statistical procedures. After discussing the dilemma of students compelled to study statistics and the accompanying domination of rote or surface learning methods, Gordon (1994a) said:

> We surely aim in our teaching to bring about learning that enables students to view mathematics as integrated and structured knowledge which enhances their understanding of the world; to encourage approaches which lead to that learning being meaningful and valued experience. (p. 17).

The development of the grounded theory of the teaching and learning of statistics commences in the next chapter. Meanwhile, in this chapter there is:

- a statement of the researcher's initial position in relation to theories of teaching and learning;
- an examination of other researchers' approaches and suggestions of how to improve the teaching and learning of statistics and how to improve education in other areas;
- the provision of a framework for the evaluation of educational innovations;
- a description of the qualitative research methodologies used in the implementation phase; and
- an examination of criteria for judging the appropriateness of qualitative research.

4.2 Positioning the teacher-researcher

One criterion for judging the adequacy of qualitative research is stated as follows:

> Assumptions are stated. Biases are expressed, and the researcher does a kind-of self analysis for personal biases and a framework analysis for theoretical biases. (Marshall, 1990, p. 193).

In the early weeks of this project, the teacher-researcher undertook a self-analysis through writing and cartooning her personal beliefs and theories regarding the nature of good and bad teaching and learning experiences, both from a teacher's and student's perspective. In this early stage, the analysis was an attempt to exorcise or set aside her own position, so that theorising about improvements to statistical education would be firmly based in data from the classroom. In the final phase of the research, further intensive self-analysis was possible. At this stage, it was also possible to examine the early ideas and to confirm them or to contrast them with those that emerged during the study. Further, it was possible to compare issues raised in the self-analysis with those raised in the literature.

In this *positioning* section there is an examination of the literature and the self-analysis of the teacher-researcher in relation to:

- conceptions of learning;
- conceptions of teaching; and
- assessment techniques.

4.2.1 Positioning: Conceptions of learning

To develop a theory of teaching statistics or improving statistical learning one must be aware of what it is that is to be accomplished. 'What is learning?' and 'what is statistical learning?' In this thesis, there was a need to clarify what *this* teacher and curriculum designer conceived of as learning. To understand how this teacher's conception of learning compares to that of other researchers it is also necessary to examine the literature for alternative conceptions.

In tailoring a teaching program to the needs of students, it is useful to address another issue, 'How do students conceive of learning?' From a positioning perspective this question was re-framed as 'What are the teacher's perceptions of student conceptions of learning?'

The teacher-researcher's view of learning

The problem that inspired this work was identified as a need to teach in a manner that induced meaningful statistical learning. The aim of teaching was to change the way that *students see the world* and in particular to enable them to view phenomena in the world from a statistical perspective. The teacher-researcher, adopted the following description of learning as representing her own:

> …the most fundamental principle ... is that learning should be seen as a qualitative change in a person's way of seeing, experiencing, understanding, conceptualising something in the real world - rather than as a quantitative change in the amount of knowledge that someone possesses....This kind of learning obviously differs from being able to remember something one has read. But is also different from 'knowing' facts and principles, and from being able

> to carry out a particular procedure. It is a change in one's understanding... From this principle of learning as a change in conceptions flow the themes and injunctions that should by now be familiar: a view of content and process in learning as parts of the same whole, an emphasis on students' conceptions and perceptions, and learning about student thinking as the key that will unlock the door to better teaching and course design. (Marton & Ramsden, 1988, p. 270).

The teacher-researcher's approach to the creation of *meaningful learning* was based on the work of Novak & Gowin (1984). The route to developing meaningful learning involved developing statistical understanding, through focussing on the rationale for *what we do as statisticians*. The unpacking of expertise and the associated identification and mapping of fundamental concepts was to bring to teaching a clarity of perspective and statistical explanation, so that the students could understand why it is that statisticians do what they do. The material was to be ordered in an appropriate manner, with linkages between concepts expounded so that conceptions of the statistical material would be both holistic and well differentiated (or detailed). In this context, the process of learning has been described as:

> Students learn statistics through: (1) assimilating new propositions into their cognitive networks, (2) altering existing propositions, (3) altering neighborhoods, or perhaps even (4) altering their networks as a whole. (Schau & Mattern, 1997, p. 171).

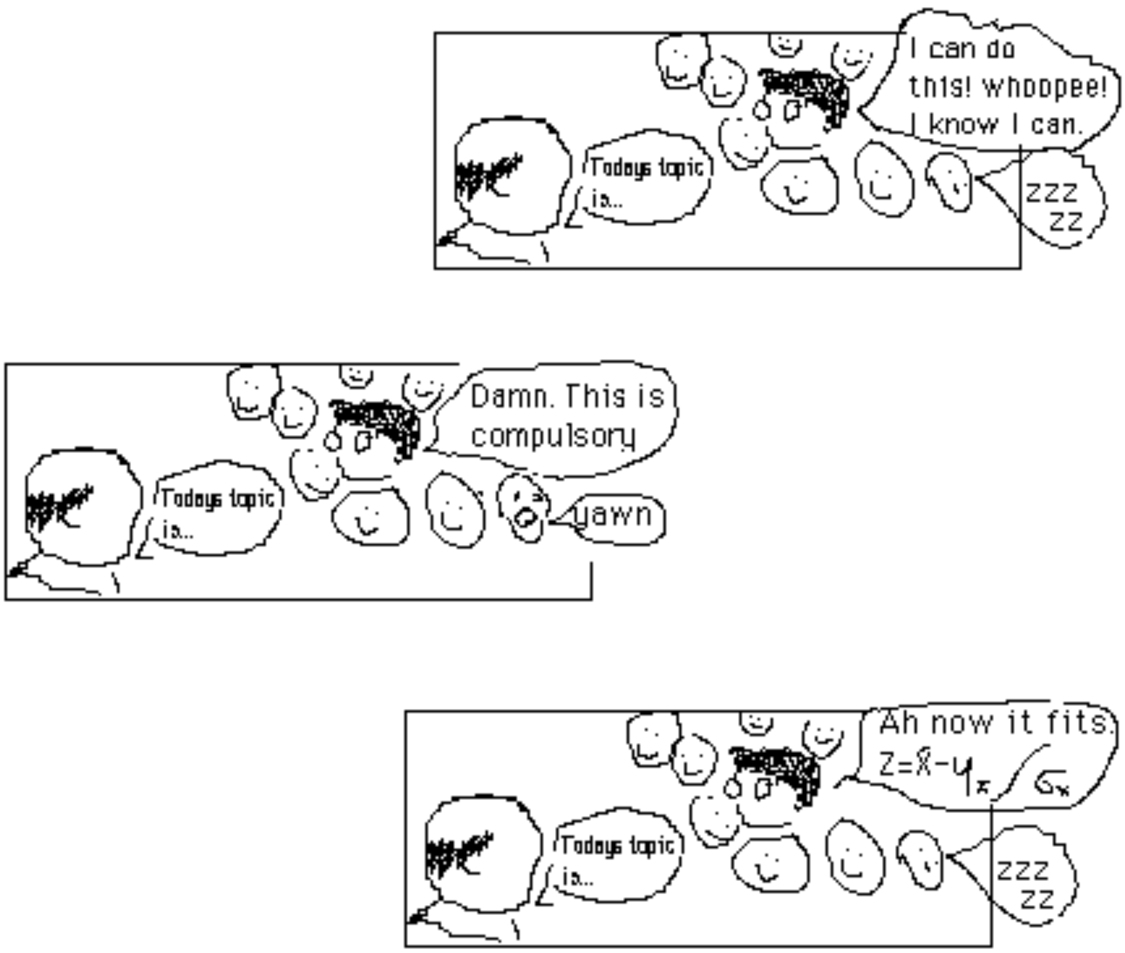
Bringing this clarity of perspective, obtained through unpacking the teacher-researcher's expertise, to the classroom does not, of itself, mean that the students have learned or will learn. In the next section, the teacher-researcher explores how it is that students are brought to this position of having learned and how it is we know they have learned.

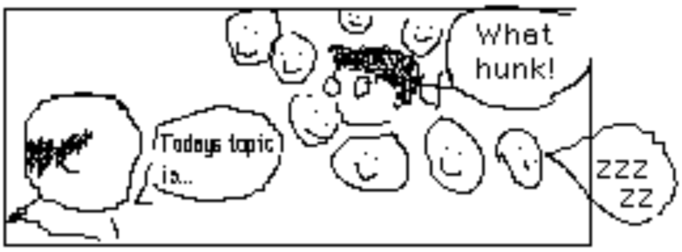
Teacher's perceptions as to students' approaches to learning

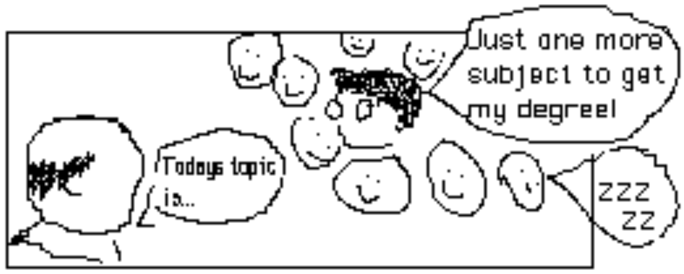
In developing an approach to teaching the teacher-researcher was influenced by her perception as to students' conceptions of learning. The teacher-researcher drew upon her many years as a student, being amongst students, her time as a teacher and working amongst teachers when she characterised students conceptions of learning or perhaps more appropriately students' approaches to and attitudes toward learning statistics.

The students taught in the implementation phase did *not want to learn statistics*. They came to University in order to learn about psychology or economics or some other discipline. The statistics subject was compulsory. From their perspective, the prospect that they could understand and enjoy statistics was often far removed. As for many of the students, who learn statistics through service statistics courses, these students anticipate that perhaps they can *learn to recognise where to apply tests and how to do the calculations* - at least until they completed or passed the subject. They anticipate that they *will not really understand statistics*. In initially cartooning this issue of students' conceptions of learning, (Figure 4.1) the teacher-researcher played with different concepts that students might hold in relation to their being a student in any given class.

Figure 4.1 Students' conceptions of the learning endeavour.







Whilst it was not necessary to identify all the conceptions at this early stage, it was important to recognise that students were likely to have varying reactions to the prospect of learning statistics. In this study, students' intentions, or perceptions of learning statistics, or indeed the strategies they used or their learning style could be identified as the statistics subject was taught. The dominant perception, based on past history, was that many of the students taught in the implementation phase were intending to study psychology and other non-statistics disciplines at University. For these students, the study of statistics was not seen to be *relevant* to their goals.

The issue of relevance was particularly noteworthy for the teacher-researcher. When students ask "How is statistics relevant?', they are recognising their anticipated roles as psychologists or economists or as professionals in some other disciplines. For example, the typical psychologist may want to help people seeking marriage guidance, addicts in withdrawal, children with learning difficulties, or people who have been traumatised. The teacher-researcher's involvement with the statistics discipline did not begin with a love of statistics, but rather through the teaching role of *helping* students who were struggling with statistics. The role of helping people has recently been identified as the motivation for the study of social sciences (Schuyten & Dekeyser, 1998). This posed a problem; how could she expect her students to experience statistics in a personally meaningful way if it was meaningful for the teacher only in the capacity of helping others. She asked in her cartooning, refer Figure 4.2, 'How is statistics relevant to a farmer's daughter?'

Figure 4.2 Statistics and the farmer's daughter.

As the cartoon also suggests, the reasons used to establish relevance may be tenuously received. As the teacher-researcher embarked on a teaching role in the statistics discipline, she needed to reframe statistics so that statistics itself would provide relevance to the students. She approached this by asking 'to what extent would her thinking alter if it was not of a statistical and mathematical nature'. In responding to this, she identified the ways in which statistical thinking was embedded in everyday life and particularly in decision making.

By slight of word (rather than hand) the term *meaningful learning* shifted. Drawing on the work of Novak and Gowin (1984), the emphasis was on a process which began with observation on events and which culminated in understanding. In this theorising, for meaningful learning to have taken place individuals must relate the new knowledge to the relevant concepts and propositions that they already know. The pedagogical approach, yet to be detailed fully, involved the identification and clarification of statistical concepts and the linkages between concepts. However, meaningful learning took on another meaning, that of *personally meaningful learning*. Later, this aspect of meaningful learning was characterised as *authenticity* (refer chapter 5). Whilst a student may understand the statistical material in a deep and meaningful way, it does not mean that the learning is salient or that it provides the learner with a sense of authenticity. In the implementation phase, in order to create a situation where learning was personally meaningful, there was a concerted effort to establish the relevance of the statistical concepts identified to everyday life and everyday thought. In learning statistics, the theme ‘we *are* statistical and numerate thinkers’ was maintained throughout all implementations. At this stage, personally meaningful learning was not distinguished from *relevance*.

The issue of relevance was highlighted by Gordon (1994a), who examined factors such as type of motivation, mathematical skills, metacognitive goals, personality and social interactions in relation to students' approaches to learning. She found that students who were compelled to undertake studies in statistics sometimes became pragmatic in their approach to learning. Goals such as the desire to pass the exams had an impact on learning, were seen to shift learning from the *slower deeper kind*, involving understanding and often requiring the restructuring of schemata, to that of *surface learning*, which is quicker. In the words of a student from Gordon's study, the way to pass exams was, '*practice a test and repeat it*, *compare the test question to the tutorial question and apply the same formula*.'

Improving Learning

Gordon's paper sparked an awareness of another body of literature regarding making improvements to learning; to this the researcher returned in the final phase of writing the thesis. Marton and Ramsden (1988) suggested that thinking in regard to the improvement of learning has been dominated by a psychological perspective and in particular through an emphasis on:

- learning strategies. These may be considered as tactics or a set of procedures for accomplishing learning, including decisions to implement skills and to plan. (Schmeck, 1988) or 'any behaviours or thoughts that facilitate encoding in such a way that knowledge integration and retrieval are enhanced...learning strategies include 'actively rehearsing, summarizing, paraphrasing, imaging, elaborating and outlining. (Weinstein, 1988, p. 291). Learners with holist strategies work with 'global predicates' and 'relations between topics' while serialists prefer to learn step by step without considering relations. (Pask, 1988).

- learning styles. This may be considered to be 'a disposition to adopt a particular learning strategy' (Pask, 1988) and examples include surface and deep approaches to learning (Marton & Saljo, 1976), The *surface* approach to learning was recently described as one whereby learners attempt to *satisfy minimum task requirements*. In contrast to deep learners as seekers of *personally meaningful learning*. (Bowden & Marton, 1998). A third category, *the strategic learner*, has also been suggested. This learner 'seeks high grades by being alert to cues from lecturers as well as being organised' (Entwistle & Ramsden, 1983). Another classification of deep-holistic and surface-atomistic draws together styles (Ramsden, Beswick, & Bowden, 1987). Reviewing the work of others Torrance and Rockenstein (1988) identified twenty-three learning styles including amongst others the right-brain-left brain, verbal-visual and logical-intuitive descriptions of traits.

- metacognitive knowledge about learning. This included 'being aware of one's own cognitive and affective states, and controlling and monitoring one's own cognitive processes, are the defining attributes of metacognition' (Biggs, 1988, p.187). Two approaches have been developed to teach students to become metacognitive and deeper learners. These involve either incorporating learning skills into the teaching of subject matter or through study skills courses (see for example Ramsden et al., 1987).

Prior to the implementation phase, and in the context of developing the statistics program, these conceptions were not consciously salient to the teacher-researcher. Certainly there was no attempt to classify students according to their approach to learning. The teacher-researcher's previous attempts to improve teaching and learning through relating student performance to psychological factors such as the *fear of success* and *need for achievement*, and to students' prior academic experience, had not resulted in improved teaching or learning practices. However it is evident that the teacher-researcher wanted to engender a deep appreciation of statistical ideas and this was embedded in her positioning statements. By way of contrast, at the commencement of teaching there was a keen awareness of appealing to students' visual, auditory and tactile senses. This awareness was based not just on students' preferred mode of attending to information (see for example, Bell, 1998) but through the presence of profoundly deaf and visually impaired students. The teacher-researcher based her teaching on an individual differences approach to teaching, using multiple teaching strategies to appeal to the different needs of student learners.

Conceptions of learning identified by other theorists.

To provide a template for comparison with this study the teacher-researcher examined the literature, addressing the question: 'what is learning?' Various authors broach this topic; see for example Marton, Dall'Alba, & Beaty (1993) and Saljo, (1979a).

One perspective has learning conceptualised in six qualitatively different ways:

(A) one's knowledge
(B) memorizing and reproducing
(C) applying
(D) understanding
(E) seeing something in a different way
(F) changing as a person. (Marton et al., 1993, p.284).

When reviewing conceptions of learning, Gerber and Boulton-Lewis, presented six conceptions, identified by Bruce & Gerber (1995), that university lecturers held of student learning. These were:

1. Acquiring knowledge through the use of study skills in the preparation of assessment tasks.
2. The absorption of new knowledge and being able to explain and apply it.
3. The ability of thinking skills and the ability to reason.
4. Developing the competencies of beginning professionals.
5. Changing personal attitudes, beliefs or behaviours in responding to different phenomena.
6. A participative pedagogic experience. (Gerber & Boulton-Lewis, 1998, p. 42).

As the statistics program was taught in a University environment, it was useful to be aware of how different lecturers conceive of learning. It is these conceptions that determine the substance of what is deemed, according to the colleagues, to be appropriate learning as judged by various assessment techniques. Colleagues will informally evaluate the students and subject according to their conception of 'what it is to have learned statistics'.

While the teacher-researcher's primary orientation was the development of statistical thinking and reasoning skills, it was clear from her teaching practices that multiple conceptions influenced perceptions as to the nature of learning and the associated teaching. Drawing upon description provided by Marton et al, the teacher-researcher's focus when speaking of learning was at the level of applying, understanding and seeing something in a different way.

4.2.2 Positioning: Conceptions of teaching

In endeavouring to teach, the teacher-researcher responded in part to what she perceived as required in the act of learning, and in part to her perceptions as to students conceptions of the statistics learning environment. In positioning herself in relation to teaching, the teacher-researcher conducted a self-analysis as to the nature of 'good and bad teaching.' She addressed questions such as 'What techniques did she use as a teacher?, and why?', 'What had she found to work?' and, 'What was she trying to accomplish?' In writing on these questions, she identified many different conceptions, which she held simultaneously, as to what it is she does as a teacher.

Self analysis: Good and bad teaching

The writings and cartoons of the teacher-researcher regarding her conceptions of teaching focussed on:

- creating an understanding of fundamental statistical concepts, that is the 'why' associated with statistical activities;

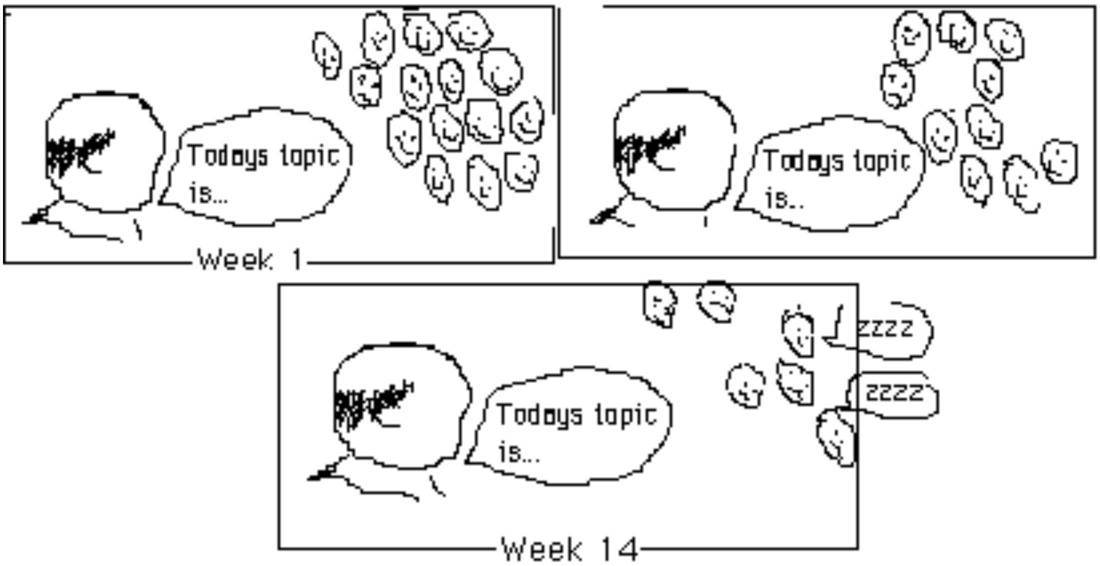
- facilitating learning;
- creating a learning environment;
- the teaching process; and
- teaching strategies or techniques.

Creating an understanding of fundamental statistical concepts. The anticipated aim of teaching involved structuring the learning experiences of students so that they could understand the conceptual basis of statistics. In so doing they were to be equipped to build their own statistical theory. This aim encompassed the ideas in the earlier section, whereby students were to become more like the expert in their knowledge structure. In terms of Gowin's Vee heuristic they would be equipped with the 'why' component in addition to the 'how' as they learned the practices of statisticians.

Facilitating learning. Associated with the concept of assisting students to understand fundamental statistical concepts was the concept of facilitating students' learning. By providing a learning environment where students could experience statistical thinking, the emphasis in teaching was on drawing out of students that for which they were capable.

Creating the learning environment. The concept cartooned (refer Figure 4.3) reflected the teacher-researcher's concern with the potential to bore through regularity of approach.

Figure 4.3 The boring teacher.



Associated with the concept of facilitating learning was the need to create an environment that was rich, varied and engaging. The objective of creating an environment, where students' own thinking was a source of surprise, developed throughout the implementation phase. This approach was in accord with the sentiments expressed by another when observing students learn:

> Most students...seemed surprised and delighted to learn that when they "created their own"... methods, they actually discovered for themselves the "proven methods." (Dietz, 1993, p. 107).

The teaching process. The teacher-as-researcher was also aware of alternative conceptions of the teaching process. One conception was that of teaching being a series of lectures and tutorials or laboratories wherein statistics are demonstrated through the analysis of real or fabricated data. As can be identified in the implementation phase, this was not an approach that was adopted and as such was rejected as a bad approach to teaching. Although in recent years as the participant-researcher's lectures to large groups have become activity based, this perception has changed somewhat. The concept, *creating a learning environment* with an emphasis on *engaging the student in learning*, was further elaborated by the adult learning theory notion that engagement is more likely to occur through active rather than passive learning. This shift in emphasis from passive to active learning can be readily identified in the literature on statistical education (see for example, ASA/MAA Joint Curriculum Committee report cited; Moore, 1997). The teacher-researcher's adopted teaching practice may be described as *activity based* with active physical and mental participation in exercises required by students. When she focussed on teaching as the use of strategies for enhancing the learning process, the teacher-researcher drew upon both traditional lecturing and activity based approaches to the *teaching process*.

Teaching strategies. As teaching was contemplated, the teacher-researcher identified several techniques which formed part of her repertoire for accomplishing various teaching and learning outcomes:

- *Linking of lectures*. The linking from lecture to the next was an important consideration when teaching. The linking was more than the identification of 'where we were up to last lecture', it provided an opportunity to re-establish the concepts previously met, and to situate the concepts in some structure of knowledge being explored.

- *Timing of teaching*. Recollections of psychology and education classes years prior suggested that timing was also important when considering memory, with retention best for material presented early, then the most recent material presented, followed by poorer retention of the material presented in the middle of a lecture. The introductory minutes were used to link lectures and review the previous lectures. The middle section of the lecture was used to introduce the new conceptual material and finer detail or *how to* information. The last minutes of a lecture were also used to reaffirm the new concepts and met in the current lecture and how they connected. Whilst new material often occurred in the middle, it was re-iterated at the end and again at the beginning of new classes. Allowing for attention spans of approximately twenty minutes the teacher-researcher, also sought to change the teaching activity every twenty minutes. The intent was to create a *new beginning* through a change in technique, style or activity.

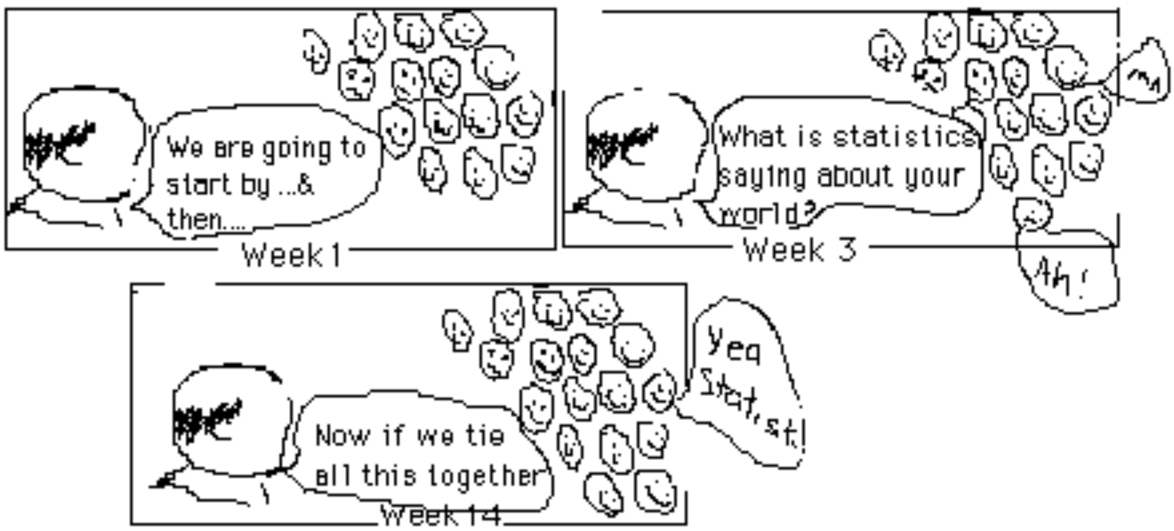
- *Use of language.* The teaching of statistics used everyday language and analogies to introduce statistical concepts and processes. Ideas generated by students as they completed assigned activities were later redefined in appropriate jargon.

- *Acceptance*. One further aspect of good teaching was deemed to be the ability to accept students. One aspect of this involved the acceptance and use of all student responses. Student answers whilst *wrong* can be used to clarify confusion between different concepts, and furthermore that which is wrong can be located in the teacher's explanation. These incorrect uses of concepts provide invaluable insight into the students' constructions of knowledge and thus are useful in determining the direction of teaching. A second aspect of acceptance involved, accepting students reactions to the learning environment as real (not negative) and as of concern to them. There was an attempt to at all time to maintain student dignity.

- *Supplementary reading*. Readings were also deemed necessary. These were necessary to provide additional detail and alternative expression of information. Formal tuition often involves the presence and often absence of students from the lecturing/tutoring process and hence there was a perceived need for supplementary reading. It was reasoned that any move away from the traditional approach to teaching statistics would also involve the preparation of supplementary reading, lecture notes and handouts. Whilst supplementary readings were identified as *necessary* for the reasons mentioned, students were not provided with a text or recommended readings. Reading materials supplied were in relation to the activities worked and the debriefing in relation to the ideas that emerged in class. The teacher-researcher remained acutely conscious that this ran contrary to current practice. In parallel strands librarians were teaching these same students how to access and use library resources. At the conclusion of the implementations, there remained an ambivalence towards providing texts and readings. The teacher-researcher remained concerned that reading about ideas, rather than generating them through active participation, would dilute the impact of those ideas. This issue is returned to when discussing the outcomes of the implementations.

- *Linking statistics to the students' world*. The students who were to be taught in the implementation phase were not committed to any particular academic program. Hence the relevance of statistics, needed to be established in terms of generic skills rather than the students' disciplines. As indicated in Figure 4.4, establishing the relevance of statistics by teacher talk, involved linking the practice of statistics to the students' everyday activities.

Figure 4.4 Linking statistical ideas to the world around us.



Teaching is a multifaceted activity. For the teacher-researcher, the provision of a rich and varied environment was seen as of primary importance. Within this environment were embedded multiple approaches to examining the major statistical concepts. The responses to all manner of interactions in the classroom were purposeful: guiding, shaping, explaining, extending, clarifying, testing, remaining patient and taking time out to observe students.

4.2.3 Other theorists: Concepts of teaching

Other researchers have also investigated conceptions of teaching. One of these elicited concepts using three techniques, concept mapping (CM), flexigrid (a form of repertory grid analysis) (F) and from observations made of a videotaped lecture (O) (Morine-

Dershimer et al., 1992). Using this approach Morine et al identified three clusters of ideas relating to:

- Curriculum. The curriculum category was composed of three subcategories:

    (1) *content/materials*. The concepts identified from the various methods were: CM- follows curriculum, uses visual aids, F- good curriculum versus bad, discipline taught, 0- related ideas to present day situations.
    (2) *student outcome*. The concepts identified from the various methods were: CM- enhance students thinking skills, promote group cooperation, teach to work independently, F- encourages creativity versus discourages ideas, stimulates thought versus does not stimulate thought, O students cannot come to conclusions on their own (-), students are not afraid to speak out if they disagreed(+).
    (3) *planning*. Only CM yielded concepts. These include variety in lesson plans and prepared for each class.

- Instruction. This was composed of four categories.

    (1) *instructional process*. The concepts identified from the various methods were: CM- asks thought provoking questions, differentiates methods to meet students' needs, follows up assignments with reinforcement activities, varies teaching styles and activities, F- answers questions versus ignores students, conveys information clearly versus does not convey information clearly, adapts to students' level versus does not, various methods versus patterned, O- questions were thought provoking(+), pushes students for more than their original response(+), examples were anecdotal and effective(+), writes key phrases on the board(+)
    (2) *classroom management*. The concepts identified from the various methods were: CM- establishes consistent guidelines, maintains order during class, good time management, F- organized versus unorganised, efficient versus inefficient, control of class versus lack of control, good pace versus too fast, O- keeps the class moving(+), pace was too fast (-).
    (3) *lesson structure*. The concepts identified from the various methods were: O- great introduction (+), might have stated goals more clearly at the beginning (-), stopped and summed up what had been said (+), drew the main conclusions rather than have students summarize (-) and
    (4) *evaluation*. The concepts identified from the various methods were: CM - assesses student understanding, good test writer, fairness and consistency in grading, O- goals clearly defined and fair evaluation.

- Social Context. This was composed of four categories.

    (1) *personal qualities*. The concepts identified were; C- patience, enthusiasm, flexibility and professionalism, F- personable versus intimidating, sense of humor versus humorless, modest versus large ego, flexible versus boring, O- confident in front of the class (+), obvious energy (+), added humor (+) and entertaining to watch (+).
    (2) *learned traits*. The concepts identified were: CM- depth of content knowledge, knowledge of different teaching methods, knowledge of students' backgrounds, F- knows material well, knows kids and observant.
    (3) *relationships*. The concepts identified were: CM- good rapport with administrators, liason between parents, faculty and students, respect for other faculty members, good rapport with students, F - friendly, respect for students, genuine concern, and good communication.
    (4) classroom role/climate. The concepts identified were: CM- teacher communicates high expectations, students share control with teacher, positive learning environment, F - expects a lot, shows personal feelings, treats all students fairly and O- abrupt (-), abrasive (-), sarcasm (-) and authoritarian presence is an effective teaching tool (+).

The teacher-reseacher's observations of teaching, her completion of a repertory grid generated similar concepts. Ramsden (1995) provided a contrasting framework when he asked the question 'What will help us, as teachers, to achieve improvement?' (p. 1). Borrowing from Bain, Ramsden compared two models of teaching and learning. One of these was deemed, *Disseminating knowledge* (DK) and the other *Making learning possible* (MLP). As Ramsden points out, and as the teacher-researcher experienced, teaching and learning may draw from both models. Characterising teaching as belonging to one or other model is a matter of emphasis. The models were compared on nine dimensions and it is useful to characterise the teaching and learning, as it occurred in the implementation phase, on each dimension.

1) *Epistemological assumptions.* In the DK model, knowledge was seen to exist separately from the people who possess it versus a perspective whereby the construction of knowledge was not deemed to be independent of the person. In this study the teacher-researcher experienced an oscillation between the two poles Moore cautions against radical constructivism, and social constructivism, particularly in the extremes:

> Taken to its limits, radical constructivism suggests that because we all construct our own knowledge, teaching is essentially impossible. No one's preconceptions can be said to be "wrong", even if the preconception in question is that the logarithm acts like a linear function. Social constructivism suggests that because knowledge is socially constructed, with no necessary correspondence to any truth "out there", teaching in any sense is akin to indoctrination. (Moore, 1997 p. 126).

In those areas of mathematical statistics which have an axiomatic basis, there is a 'truth', a knowledge which existed independently of the observer. Indeed, as Moore remarks it is difficult for any mathematician to ask "when does 2+2 not equal 4?" (Moore, 1997, p. 126). However, for the greater part of the thesis and teaching, and particularly where statistical thinking was at issue, the teacher-researcher adopted a stance wherein knowledge was not viewed independently of the observer. For the teacher-researcher, different constructions of knowledge came from the different ways in which elements of knowledge were related, or from differences in the implications that were perceived to follow knowledge atoms.

2) *Academic and social environment.* In the DK model, knowledge is learned through 'individual study and practice…other students provide competition but are otherwise marginal to learning'. In the MLP model 'effective learning occurs in an

environment that mimics social systems of inquiry; social interaction and cooperation are essential to the negotiation of understanding' (Ramsden, 1995, p.19). The academic environment reflected the oscillation in epistemological assumptions, and the mathematics versus statistics dichotomy. The dominant environment for learning mathematics (principally out of class) involved individual study and practice, although students could and did often work with others when they felt they needed assistance, whilst in the statistics learning environment the emphasis was on social interaction and cooperation.

3) *Student learning*. The DK model is teacher-focused, students 'practise procedures, produce correct answers, reproduce knowledge accurately' (Ramsden, 1995, p.19). The MLP model is learner focused, with abstract concepts and principles derived from experience. The emphasis is on students' construction of personal understanding and the emulation of experts' methods. The emphasis in 'teaching' statistics in this study on developing students thinking such that they used key statistical concepts when providing solutions to problems was clearly from the MLP model.

4) *Teacher's role*. In the DK model the teacher's role is to:

> ...ensure that the ever-expanding content is covered. Organise and present knowledge well. Arrange suitable teaching activities, including lectures, tutorials and labs as appropriate. Rely on students to understand and absorb presented knowledge and procedures.' (Ramsden, 1995, p.19).

However, the MLP model best describes the teacher's role in this study. The emphasis throughout the implementation phase involved the use of activities to generate statistical thought and dialogue. The teacher's role was to clarify, extend and assist in the restructuring of ideas. This is consistent with the MLP model where the teacher's role is to:

> ...limit content to the essential. Model the methods of practice and scholarship in the field. Design diverse tasks strongly related to learning goals. Challenge misconceptions and build understanding through dialogue. Constantly monitor student understanding and intervene whenever necessary. (Ramsden, 1995 p.19).

5) Desirable learning and assessment tasks. Desirable learning and assessment tasks. In the DK model, assessment tasks tend to be

> '...well-structured problems and standard exercises with high reliability. "Decontextualised practise": parts are studied separately and only brought together towards the end of the course. Tasks provide feedback at the end of the unit, or not at all' (Ramsden, 1995, p.19).

This would provide an apt description of the assessment for mathematics. However the learning and assessment tasks for statistics in this study would be best described according to the MLP model. For this the assessment tasks are:

> ...loosely structured problems and realistic tasks requiring student decision making. "Situated practice": tasks of increasingly complexity that incorporates essential skills and knowledge. Tasks provide continuous feedback. (Ramsden, 1995, p.19).

6) *How is teaching improved?* This refers essentially to the environment in which teachers operate, but, in commenting, the researcher is addressing what she responded to in the environment. In the DK model, teaching is improved 'mainly through practice; driven by extrinsic rewards' (Ramsden, 1995, p. 19). The route to improving statistical education in this study has not been without extrinsic rewards, however it conforms closely to the MLP model whereby intrinsic interest has been a driving force, as have 'repeated cycles of reflection and action' upon teaching.

7) *Approach to management and leadership teaching.* In the DK model the context within which teaching takes place is 'essentially transactional: assigning tasks and rewarding their successful completion', whilst in the MLP model the approach is 'essentially transformational: creating an enabling environment and pursuing a moral vision. (Ramsden, 1995, p. 19).' The learning development environment within which this teaching was conducted was an enabling environment. Teaching was clearly the singular most important role for each academic. Daily, informal discussions with other staff would centre on how to solve or approach problems encountered with teaching and learning.

8) *Evaluation and audit.* In the DK model, evaluation is 'measurement focused, externally directed and value-free. Preferred indicators are quantitative, such as pass rates and student ratings' whilst evaluation for the MLP model is 'process focused, user directed and permeated by values. Preferred indicators are qualitative, such as student comments and evidence of changes in conceptions' (Ramsden, 1995, p. 20). Whilst the dominant form of evaluation was conducted internally and 'process focussed', in this study, pass rates, marks and student ratings are also used to enable comparisons between different implementations.

9) *Educational effectiveness.* Improving education in the DK model is 'essentially technical: a problem to be solved' whilst for the MLP model it is 'essentially problematic: an enduring human dilemma. (Ramsden, 1995, p.20). As Ramsden concluded, these dilemmas are to be repeatedly encountered. Similar issues resurfaced, as the teacher-researcher moved from one class of students with one subject in this study to other subjects for other students in the reflective phase. The solution to similar problems needed to be formulated again with the particular group of students in mind.

4.2.4 Positioning: Assessment

Self-analysis: the nature and role of assessment in learning

Assessment was originally conceived of by the teacher-researcher as just another teaching strategy. However, the relationship between assessment and learning assumes a greater role in the literature than many of the other strategies identified. Ramsden, said that in improving the quality of higher education:

> University students' experiences of teaching and assessment matter more than particular teaching methods in determining the effectivenesss of their learning… And, if teaching is about helping to make learning possible, assessment becomes defined as being about understanding students and what they have learnt. Effective assessment helps students develop the skills of self-assessment. (1995, p. 4).

In the positioning writings, the teacher-researcher asked 'What impact does the nature of assessment have upon the learning of students?' At this stage she was aware of the possibility that:

> …the student goals of passing and satisfying course requirements, which are often facilitated by cookbook recipes for the selection and calculation of various statistics, may mitigate against deep learning, that is, mitigate against the learning of how statistical concepts relate and of processes are involved in statistical decision making. (Writings, 1993).

Yet there was a need to address the issue of 'How it is we know they have learned'. The awarding of University degrees is dependent on evidence that students have learned. The teacher-researcher drew heavily on the theoretical work of Novak & Gowin, (1984) and her own experience in mathematics and statistics remedial work with students. Her practice of identifying, the conceptions held by students and missing concepts, and her subsequent development of conceptual structures with students continued. Assessment was orientated to revealing students' conceptual understanding of statistics. As she considered meaningful learning, she wrote about the perspectives of other theorists:

> A further criterion for meaningful learning to have taken place is that, individuals must relate the new knowledge to the relevant concepts and propositions that they already know (Novak & Gowin, 1984). In this way meaningful learning contrasts with rote learning, which involves knowledge being arbitrarily incorporated into a person's knowledge structure without interacting with what is already there as happens when simple memorisation of new knowledge takes place.
>
> The techniques of concept mapping, semantic networks and Gowin's Vee provide mechanisms for creating visual images of both experts' and students' knowledge structures and thus enable a ready comparison between the two. A further benefit of this approach is that it may be used in conjunction with students' maps in order to identify their misconceptions whilst at the same time facilitating the acquistion of *shared meaning* between teachers and students (Novak & Gowin, 1984). (Diary, 1993)

4.3 Literature Review: Improving Statistical Education

In this section, the review, which followed the implementation of the curriculum, focuses on changes suggested by the literature as a means of improving statistical education. To facilitate discussion, the suggestions have been classified as belonging to one of three categories:

- curriculum, particularly curriculum content;
- pedagogical practices, with emphasis on the strategies adopted in teaching; and
- assessment.

Only a few of the suggestions for improving statistical education have been trialed and comprehensively evaluated. However it was thought that the examination of the suggestions for improving statistical education provided a context for reflecting upon:

- coherence between the fundamental concepts espoused by experts and suggested curriculum content;
- teaching practices proposed in the literature and those documented throughout the implementation phase (refer Chapter 5); and

- assessment practices used and those documented.

4.3.1 Curriculum / Academic Content

The term curriculum may be used to refer to the aims, objectives and academic content, but in a wider sense may also refer to teaching practices and assessment. In this section, the emphasis has been on a statement regarding possible academic content.

The first three statements regarding academic content are in many ways consistent with each other, differing in the detail rather than the sentiment. The statements may be seen to reflect collected wisdom, rather than evidence as to the efficacy of changes in curriculum as delivering better statistical understanding on the part of students. Three examples of this are provided by:

(1) the recommendations of the Joint Curriculum Committee of the American Statistical Association (ASA) and the Mathematical Association of America (MAA) is that the nature of first courses in statistics should be as follows:

1. Emphasize the elements of statistical thinking:

(a) the need for data,

(b) the importance of data production,

(c) the omnipresence of variability,

(d) the measuring and modelling of variability

2. Incorporate more data and concepts, fewer recipes and derivations. Where ever possible, automate computations and graphics. An introductory course should:

(a) rely heavily on *real* (not merely realistic) data,

(b) emphasize *statistical* concepts, e.g. causation versus association, experimental vs. observation and longitudinal vs. cross sectional studies,

(c) rely on computers rather than computational recipes,

(d) treat formal derivations as secondary in importance.

3. Foster active learning, through the following alternatives to lecturing:

(a) group problem solving and discussion,

(b) laboratory exercises,

(c) demonstrations based on class generated data

(d) written and oral presentations

(e) projects, either group or individual.

(ASA/MAA Joint Curriculum Committee cited; Moore, 1997, p. 127).

(2) Drawing on their own earlier work (Gal & Garfield, 1997), Garfield and Gal (1999) listed the big ideas that underlie statistics. In terms of Gowin's Vee the first, fourth, fifth and seventh of these points could be described as the *conceptual* component aspects, that is the *what* and *why*. The second and third points could represent the *doing* component with an additional component, the ability to communicate. The big ideas were listed:

1. Understand the Purpose and Logic of Statistical Investigations…
    - The existence of variation
    - The need to describe populations by collecting data.
    - The need to reduce raw data by noting trends and main features through summaries and displays of the data.
    - The need to study samples instead of populations and to infer from samples to populations.
    - The logic behind related sampling processes.
    - The notion of error in measurement and inference, and the need to find ways to estimate and control errors.
    - The logic behind methods (such as experiments) for determining causal processes.
2. Understand the Process of Statistical Investigations …
    - Formulating a question.
    - Planning the study (e.g. approach and overall design, sampling, choice of measurement tools).
    - Collecting and organizing data
    - Displaying, exploring, and analyzing data.
    - Interpreting findings in the light of the research question.
    - Discussing conclusions and implications from the findings, and identifying issues for further study.
3. Master Important Procedural Skills…
4. Understand Probability and Chance…
    - Concepts and words related to chance, uncertainty, and probability that appear in our everyday lives, particularly in the media.
    - It is important to understand probabilistic processes in order to better understand the likelihood of events in the world around us, as well as information in the media.
    - Probability is a measure of uncertainty
    - Developing a model and using it to simulate events is a helpful way to generate data to estimate probabilities.

• Sometimes our intuition is incorrect and can lead us to the wrong conclusion regarding probability and chance events.

5. Develop Interpretive Skills…
6. Develop ability to Communicate Statistically
7. Develop Useful Statistical Dispositions (Garfield and Gal, 1999, pp. 2-3).

(3) Nicholls (1999) draws upon Smith's (1996) work when discussing curriculum content from the context of *preparing students for the workplace.* In the workplace the statisticians' roles involve:

> …providing theory and protocols to guide and discipline all forms of quantitaive investigatory procedure. Such theory and protocols fall under the following kinds of heading:
>
> - the framing of questions;
> - design of experiments and surveys;
> - drawing up protocols for data collection and recording;
> - collection of data by sampling or observation;
> - monitoring compliance with protocols;
> - monitoring data quality, data storage summarization and presentation;
> - stochastic modelling;
> - statistical analysis;
> - model criticism and assessment of assumptions; and
> - inference reporting and the use of results for prediction, decision making and hypothesis generation. (Nicholls, 1999, p. 129).

The literature also reveals two other contrasting approaches to these statements on curricula. Often the approaches may overlap, but they may also be recognised as distinctly different approaches to statistical education. These approaches are teaching through resampling and teaching statistics from a Bayesian perspective.

Resampling

Statistics may be taught through *resampling* (Peterson, 1991; Seiter, 1994; Simon, 1994) rather than formulaic methods (Simon, 1994) found in traditional approaches to teaching. This is not inconsistent with the first of the curriculum statements provided by the ASA and MAA, when they sought to *treat formal derivations as secondary in* *importance*. The shift away from teaching statistics in a *mathematical* and *theoretical* way to teaching through an emphasis on applied statistics with an emphasis on exploratory data analysis has been well-documented. (For a fuller discussion of this historical shift refer Hawkins, Jolliffe, & Glickman, 1992).

Simon describes resampling as follows:

> Resampling refers to the subclass of statistical problems done with simulations; it includes bootstrap and Fisherian permutation methods as well as simulation techniques for dealing with a variety of other problems such as binomial proportions' (Simon, 1994, p. 290).

Resampling can be done, in principle by enumeration. But, in practice, the number of possible cases may be so large that simulation is necessary. The resampling method may be characterised as a method which:

> ...emphasises intuitive reasoning and involves using a computer to perform experiments on the available data -- a process known as resampling -- to get meaningful answers to statistical problems. In his scheme students avoid having to tangle with mysterious formulas, cryptic tables and other forms of mathematical magic to get their results (Peterson, 1991, p. 56).

The *big ideas* of statistics such *as statistics is about variation* may be demonstrated through these resampling procedures rather than deduced from theory. Simulations are also often used within traditional courses to develop an intuitive understanding of classical theory. However *resampling* may also be considered to be a pedagogical approach to learning statistics, as the approach has implications for the inclusion or more specifically *non inclusion* of theory and formula that traditionally reside in introductory courses. The resampling approach may be directed more to solving statistical problems or decision making rather than to generating an appreciation of classical theory.

The Bayesian Perspective

The second contrasting approach is teaching statistics from a *Bayesian* perspective (see for example, Blackwell, 1969) rather than a *frequentist* perspective that underlies most traditional statistics courses (Berry, 1997). The approach was described as follows:

> Bayesians take a larger view, and one not limited to data analysis. In particular, the Bayesian approach is subjective, and requires assessing prior probabilities. This requirement forces users to relate current experimental evidence to other available information… One basic idea, of course is Bayes' theorem. Another is that all uncertainty is represented using probability. Learning the basics of Bayesian statistics requires following logical thought processes (Berry, 1997, p. 241.)

Teaching from a Bayesian perspective may also be undertaken using resampling methods (Albert, 1993; Smith, 1996; Smith & Gelfand, 1992) or can be taught from theoretical, formulaic methods. Within all three perspectives there can be the orientation to make statistics more or *less* mathematical (see for example, Higgins, 1999).

Examining further these less recognised approaches to teaching statistics, is beyond the scope of this thesis. However in dismissing the approach in this manner one needs to be mindful of the comments made by others:

> I contend that the Bayesian introductory statistics courses…should be considered acceptable alternatives to the traditional frequentist approaches. (Short, 1997, p.263).

In making this statement, Short illustrates that the approach to teaching, just as the earlier approach recommended by the ASA/MAA Joint Curriculum Committee, may emphasize statistical thinking; involve more data and concepts; involve less theory and recipes and foster active learning.

One can also approach the improvement of statistical education from the perspective of changing pedagogical practice rather than curriculum content. Indeed some of the specifications listed under curriculum content make reference to the manner in which statistics is taught.

4.3.2 Changes in pedagogical practice

In order to evaluate improvements or otherwise in statistical education associated with a change in curriculum or pedagogical practice one needs to consider:

- the nature of the outcomes improved;
- the context for making the improvement;

- how the agent of change is identified and described;
- the quality of the evidence; and
- the evidence itself.

Outcomes defining better education

When authors document a change in teaching technique, they tend to identify some changed outcome. Often there is an implication that the change produces better statistical education. However, there are varied outcomes of interest to educators. The innovations need to be considered in terms of the associated educational outcomes. Typical terms in studies regarding improvements for statistical education include changing or influencing motivation, engagement, understanding or analytical thinking skills, attitudes such as anxiety, fear or interest, the retention of concepts; and the amount learned.

To look at changed learning outcomes in greater depth one could draw on a number of educational taxonomies (or learning hierarchies) when examining the nature of learning (see for example, Biggs & Collis, 1982; Gagne, 1977). If one were to use *Bloom's Taxonomy of Educational Objectives* (1956), one would classify the students' type of learning in the following hierarchy:

- recall knowledge, remember facts,
- comprehend (or be able to explain in ones own word), understands relations and context;
- apply facts, rules or procedures to new areas;
- analyse, separating the whole into parts, to classify, compare and contrast;
- synthesise or combine ideas to form a new whole;
- elucidate, develop judgement, or establish which criteria they would use to make a choice.

These earlier taxonomies have been extended to include structural and higher-order cognitive, metacognitive and motivational learning outcomes. These extensions recognise the different learning environments where knowledge may be either well or ill-structured. The extension to the taxonomies was done in the context of:

> ...'real-world(situated) problems' emanating from ill-structured knowledge domains . To solve these problems students must be able to '(1) define the problem, (2) decompose the problem by relating it to domain knowledge, (3) hypothesize alternative solutions, and (4) evaluate the viability of those solutions'. (Jonassen & Tessmer, 1996/7, p. 32).

The skills associated with completing these tasks include (1) ampliative skills[1], such as 'drawing analogies, making inferences and constructing arguments', (2) self-knowledge including awareness of learning styles, learning strengths and level of knowledge, (3) executive control or metacognitive strategies, and (4) conation (or motivation). Executive control strategies for planning include:

> (a) assessing/estimating task difficulty, (b) setting learning goals, (c) selecting or determining strategies for accomplishing tasks, (d) allocating cognitive resources, (e) assessing prior knowledge (also a part of self knowledge), (f) assessing progress toward the goal, and (g) checking your performance for errors. (Jonassen & Tessmer, 1996/7, p. 33).

One could also consider as outcomes of teaching examining changes in students' critical thinking. Students should be able to understand, formulate, analyse and assess problems or questions. This they should be able to do in relation to the purpose or goal of thinking, the frame of reference, assumptions made, central concepts and ideas involved, principles and theories used, evidence, reasons or data advanced, claims made, inferences, reason and formulate thoughts and the implications and consequences that follow. (Paul, 1992).

In statistical education research few studies examine outcomes in terms a taxonomy of learning. An exception to this is a study that examines student understanding in cohorts of students over different time periods and grade levels. Specifically the study examined student responses in terms of:

---

[1] Ampiative skills are considered as 'distinct from problems solving skills because they are not directed at arriving at a solution as much as extending the learner's knowledge of a domain'. (Jonassen and Tessmer, 1996/7, p.32).

> (1) a basic understanding of statistical terminology; (2) an understanding of statistical language and concepts when they are embedded in the context of wider social discussion; and (3) a questioning attitude which can apply more sophisticated concepts to contradict claims made without proper statistical foundation. (Watson, 1998, p. 793).

Whilst the study was not a comparative study in the sense of examining improvements due to changes in pedagogical practice, it could be used to provide a measure of understanding, and hence be of use to curriculum planners. Watson drew upon the work of Biggs and Collis (1982) who proposed the SOLO *(Structure of the Observed Learning Outcome*) taxonomy involving the levels:

- Prestructural - irrelevant response;
- Unistructural - the use of one obvious piece of given data;
- Multistructural - the sequential use of one obvious piece of given data;
- Relational - the integration of the given data to form a unique conclusion or generalization; and
- Extended Abstract - the use of multiple interacting abstract systems to form a response. This may include forming a general hypothesis, assessing the quality of models and accepting open-ended answers. (Kjollerstrom & Martensson, 1999, p.18.).

Assessing the outcomes of statistical learning has involved the use of the classification of *the structural elements of meaning in mathematical objects*. These structural elements include:

> a) intentional elements: concept definitions, propositions, procedural descriptions; b) Extensional elements: situation-problems, exercises, tasks; c) Notational (or instrumental) elements: Systems of symbols, computing and representational resources (tables, graphs, texts, etc.)(Godino, 1998, p.905).

Some researchers structure their teaching according to the PCAI cycle (i) **P**ose the question and produce a hypothesis (ii) **C**ollect the data (iii) **A**nalyse the results and (iv) **I**nterpret the results (see for example (Ben-Zvi, 1997)). When teaching is structured in this manner it would be possible to examine critical thinking. Ben-Zvi's study was not a comparative study. However, he suggested based on observational techniques identifying students' uncritical thinking, the meaningful use of representations, the meaningful handling of multiple representations and creative thinking, that the meaningful learning of statistics was promoted through this form of teaching.

One problem found when reviewing the papers on improving statistical education is that when investigating the nature of learning in statistical education *the nature of learning is taken for granted* (Saljo, 1979, p.446). To improve the quality of learning one would adopt teaching practices that improved the higher order levels of thinking and it is therefore these levels of learning that would need to be assessed.

Context for evaluation

When evaluating the outcomes of changes to teaching one must also consider the context of that learning. That which improves learning for students involved in distance education programs may not have the same impact when applied in on campus programs. For example, whilst distance students may have a positive reaction to edited videos of lectures (Godden, 1998) or interactive learning environments (Portier & Van Buren, 1995), the same might not be found for their on-campus counterparts.

Describing the agent of change

Some authors describe their work in broader theoretical or epistemological terms, for example they are moving from a passive learning to an active learning environment (see for example, Scheaffer, 1994; Shaughnessy, 1977), using a problem-based learning approach (Ryan, 1998; Wang & Wong, 1998) or say to adopting a constructivist perspective. For example, Moore's (1997) *commentary* described a general shift in pedagogy from 'the "information transfer" to the "constructivist" view of learning' and the associated shift from listening and reading to active participation. Student learning, by implication was improved by:

> Group work in and out of class; explaining and communicating; frequent rapid feedback; work on problem formulation and open-ended problems. (p. 125).

Moore could be describing what in Ramsden's (1995) terms could be viewed as a change where students learned in a social environment. In many such commentaries, where authors describe their work in terms of the learning environment or epistemological assumptions, the inference is that there may be multiple changes to both teaching and assessment practices.

Some articles examined make multiple attributions as to the strategies used to improve statistical teaching and learning. Two different examples of these are as follows:

- McBride (1996) provides an example where multiple strategies are used to induce analytical thinking skills, as measured by the Whimbey Analytical Skills Inventory. McBride oriented teaching toward active learning, used a range of strategies. These included students working in pairs, performance based exams involving the analysis of many data sets, use of statistical software, open textbook exams. Using a pre-test, post-test design, without a control group, he concluded that students given the opportunity to engage in active learning were more willing to tackle questions at the end of session and that they had improved their general analytic skills.

- Selvanathan & Selvanathan (1998) provided another example of using multiple strategies to make the teaching of statistics a success. They described the key characteristics of good teacher performance and the methods of teaching. Performance of the teacher involved planning, implementation and evaluation.

Many of the articles discussed in this section make few attributions regarding the cause for change in outcomes. One of the difficulties associated with classifying any attempt to improve statistical education is that whilst the author may indicate one focus, perhaps one change in pedagogical practice, there may be indications of other pedagogical changes or practices which may have effected outcomes. Indeed for all these studies one must remain somewhat sceptical as to what has actually changed as teaching and learning environments are incredibly complex and dynamic.

Quality of the evidence

Most articles reviewed in this section *do not* provide comprehensive evidence or evaluations regarding the outcomes of the suggested educational innovation in pedagogical practice. This difficulty is described as follows:

> ...little is know about or has been published on the methodology of statistical education research. As is the case in social and medical research, it is often not practical to

> experiment in educational research and it is not always ethical to do so. Controlled experiments are particularly difficult to do. (Jolliffe, 1998, p. 804).

However, articles that merely provide a commentary have been accepted in this study as being based in the teacher-author's experience of statistical education. The articles have been classified in terms of those which provide a *commentary,* that is no formal evaluation, and those which *to some extent* might be considered in the words of Hawkins to provide an 'evidence-based understanding about the teaching/learning process' (p. 144). As the evidence provided is generally poor, emphasis in this section is on identifying the multiple conceptions as to what has changed and the attendant impact on statistical education, rather than whether or not correct attributions have been made regarding the cause of the change.

Evidence and Commentary

To illustrate the range of approaches used to effect an improvement in statistical education a selection of articles, drawn primarily from *the Fifth International Conference on Teaching Statistics* (1998) have been reviewed. They have been classified in Table 4.1, according to the key descriptor of the suggested pedagogical change, the learning outcome and whether or not the outcomes, were supported by commentary or evidence. The evidence supplied is generally weak. The purpose of the review is therefore re-iterated as simply illustrating a range of strategies suggested for improving statistical education and the type of evidence associated with the claim. Studies have been classified according to the categories of:

- *activity based learning*. The implementation of activities has been directed toward improving a range of outcomes, conceptual understanding, engagement, motivation and the reduction of the fear associated with learning statistics. All five studies included in Table 4.1 provide commentaries as to the nature of activities that can be used to improve learning.

- *using and/or collecting real data.* Many accounts focus on the nature of data that is supplied or collected by students. For example, Roberts (1993) described a shift in student learning with a change in teaching. Teaching moved from the traditional approach of a theory lecture followed by practical exercises, to approaches increasingly using real data sets and finally to an approach wherein students generated their own data, this being followed by lectures with the relevant theory. Students, according to Roberts were considered to be more *engaged* in learning when they generated their own data. Other outcomes as described by these studies were creating understanding, motivation and interest.

- *using technology, changing computer software, using multimedia or the internet.* One of the difficulties in attributing the success or otherwise of technology for improving learning or attitudes toward learning is that the implementations themselves need to be good implementations in order to adequately assess the outcomes. Poor technology usage should lead to poor learning outcomes and such some studies may not provide for an appropriate evaluation of the potential for technology to improve learning. Software packages used as an adjunct to teaching may vary in the embedded paradigm; instructional (drill and practice), revelatory (learning by discovery), conjectural (testing ideas and hypotheses) and emancipatory (reducing the number of calculations) (Ng & Nobar, 1998). In reviewing these articles, actual evaluation of whether or not the technology has been appropriately used is beyond the scope of the thesis. Implementations can involve comprehensive learning statistics systems such as StatPlay (Cumming & Thomason, 1998) and the associated 'Play it again SAM facility' (Les, Maillardet, & Cumming, 1998), Models 'N Data (Stirling, 1995) and hypermedia (Ciaccio, 1998). Implementations also involved the use of graphic calculators (Bilotti-Aliaga, 1998; Lipson, 1998; Situmeang, 1998) and generic spreadsheets (Carter & Mougeot, 1998). In a meta-analysis of articles from 1967-1997 Fitzgerald (1998) concluded that students performed as well or better when exposed to computer-assisted instruction, but cautions that the finding is not homogeneous. As with many other studies, there is insufficient information to further examine Fitzgerald's findings. Fifteen of the studies included in Table 4.1 report various uses of the use of technology, graphics calculators, computers, the internet and particular forms of software, such as spreadsheets or Minitab. These do not reflect favourably on the use of technology for improving learning although providing comparable learning environments to the traditional seems plausible.

- *selecting appropriate textbooks.* This is of particular interest in the context of the earlier review of how the term variability and variation often fails to be made explicit in statistics textbooks. The authors of these papers make commentary as to what they consider appropriate in textbooks in order for them to facilitate learning or some other outcome such as decreasing anxiety.

- *thematic approach.* Another approach to improving the teaching and learning of statistics has been to use a thematic approach. Teaching statistics may take place in the context of a theme, such as total quality management (TQM). Some use the idea of multiple themes within subjects or an overall subject theme and in each case the outcome has been described in terms of increased motivation, interest or engagement.

- *miscellaneous*. There is a host of other strategies, aimed at changing some one or other educational outcome, whether it is affect, (the emotional response associated with the experience) or amount learned, retained or understood. Strategies include co-operative learning, the use of workbooks, video, the use of case studies.

Table 4.1 Improving Statistical Education - commentaries and evidence

Innovation and Author	**Outcome**	**Evidence (even though often poor) or Commentary (no formal evaluation)**
Strategy: Activities		
Activities to reveal principles of experimental design. (Cobb & Miao, 1998)	Engagement, motivation	*Commentary*: The activities used to examine experimental design are discussed. The authors conclude that the activity is effective in creating discussion, engaging and revealing design concepts.
Humorous and interactive group activities (Lackey, 1998)	Removes fear of learning, easier to retain concepts	*Commentary*: The article focuses on a description of fun activities.
Activity to reveal design principles (Ngee, 1998)	Motivation	*Commentary*: Students are eager to be involved in 'planning, brain storming, data collection and analysis." p. 222.

Innovation and Author	Outcome	Evidence (even though often poor) or Commentary (no formal evaluation)
Workshop statistics (active learning) (Rossman & Chance, 1998)	Conceptual understanding	*Commentary*: The article provides a description of activities and implementation of the workshops.
Active Learning strategies (Witmer, 1998)	Understanding	*Commentary*: The author discusses the nature of activities that can be used.
Strategy: Real Data		
Working with real data. (Burrill, 1991)	Understanding	*Commentary*: The article discusses implementation issues.
Case studies of realistic economic problems. (also teams of two, Minitab, laboratory & lectures & report writing) (Carlson, 1999)	Motivation	*Evidence*: • 80-90% Students responded favourably to the use of cases. • a correlation was found between the projects and examination marks. Commentary: suggested that understanding and motivation had improved.
Hands on involvement with data (Rossman, 1994)	More enjoyable and productive learning environment	*Commentary*: The article describes the types of activities that can be used.
Use real world data. (Singer & Willett, 1990)	Motivation, interest and understanding.	*Commentary*: The focus is on selecting or finding data sets. The approach reveals not only *how* data are analysed, but *why* they are analysed (p. 223.) These data need to be • authentic, raw data, actual measurements • with background information re purpose & design • interesting, controversial, topical, historical, • allow multiple analyses • result in substantial knowledge about the world.
Student-generated data (Stedman, 1993)	To engender and maintain interest and hence improve understanding	*Commentary*: Anecdotal evidence suggests improved understanding.
'Decision making context with client driven problems; computer-based and with a conceptual umbrella over assignments' (Swanson & McKibben, p. 159)	To increase learning and retention	*Evidence*: • The course receives high evaluations • 'informal evidence suggests students retain a "structural" picture of statistics that allows them to quickly regain a good understanding of individual aspects of statistics' p. 164.

Innovation and Author	**Outcome**	**Evidence (even though often poor) or Commentary (no formal evaluation)**
Strategy: Software Technology and multimedia		
Excel (Carter & Mougeot, 1998)	Deeper understanding	*Commentary:* Students got a better 'feel' for the data and statistical analyses overcoming problems that may be encountered by the 'unprepared use of traditional "black box" type statistical packages.' (p. 847).
Internet versus internet plus laboratory Differences were attributed to access to the instructor and to the greater access to cooperation with fellow students in the on campus group. (Chizmar & Walbert, 1999)	Amount learned	*Evidence*: This was gained from formative feedback and one minute papers, eg. • Work is described in the context of Chickering & Gamson's (1987) seven principles for good practice: engaging students in student-student and student-faculty contact, active learning techniques, providing prompt feedback, developing reciprocity and cooperation among students, emphasing time on task and respecting diverse talents and ways of learning. • Outcomes measured by the number of pre- and post-tested questions answered correctly. The internet group did *not* improve as much as the on campus laboratory group. • The on-campus group also had higher final grades than the internet group.
DSTATS - descriptive statistical system (Cicchitelli, 1998)	Understanding of statistical concepts.	*Evidence*: A group using textbook was compared with those using DSTATS on topics from descriptive statistics and found those students using DSTATS did better on a 85 item test than 70% of students using the textbook approach.
Computer packages (Dallal, 1990)	Motivation	*Commentary*: The authors describe how the computer was used to establish the *how* of statistics complementing learning *why* and *when.* 'Teaching portions of statistical computer packages can give students an appreciation of the knowledge and care that must go into using techniques' (p. 265).
Computer Based Learning Materials (Davies, Lees, Smith, & O'Neill,)	Learning Pace of learning Attitude	*Evidence*: Student related outcomes were as follows: '(7) Students using CBL materials exhibited neither more nor less learning gain compared to students studying by conventional means; (8) found the materials easy to use and were positive about the use of the computers for learning; (9) With a higher ability were able to work through the materials at a faster pace' (p. 222).

Innovation and Author	Outcome	Evidence (even though often poor) or Commentary (no formal evaluation)
Computer micro-world.(delMas, Garfield, & Chance, 1998)	Understanding	*Evidence:* The results of a comparative study found that: • 'an activity which asked students to test their predictions and confront their misconceptions was found to be more effective than one based on guided discovery...while computer software can provide the means for a rich classroom experience, computer simulations alone do not guarantee conceptual change'. p 1083.
Visual representation through StatPlay (Finch & Cumming, 1998)	Conceptual change, Overcoming misconceptions	*Evidence*: These authors suggest 'Well designed visual representations can help in understanding sampling variability. Visual representations can also be confusing.' (p. 901).
Visualisation and simulation of concepts (Finzer & Erickson, 1998)	Construct their own understanding of concepts	*Commentary*: The author discusses the tools of dynamic dragging, visualisation, simulation, and networked collaboration as a means of creating statistical ideas.
Simulations (Kennedy, Olinsky, & Schumacher,)	Clarify theoretical methods	*Commentary*: Simulations can be used to clarify theoretical methodology.
Computer-aided instruction (Lee, 1998b)	More knowledge	*Evidence*: • Completed a needs analysis to identify the gap between students current and desired level of learning • Defined teaching functions to be remediated • reviewed objectives and preliminary knowledge • A pre-test - post-test showed improvement in knowledge (as to be hoped for any teaching).
Computer graphics and simulation supplementing data analysis. (Active learning) (Marasinghe, Meeker, Cook, & Shin, 1996)	Understanding through experience	*Commentary* : This author describes the modules used.
Electronic learning environment. (Schuyten & Dekeyser, 1998)	Learning outcomes	*Evidence*: Comparing an electronic based independent learning system, paper based system and traditional lectures the following were found: • no statistical differences were found in study outcomes; • lectures were preferred for reasons of social contact; but • the electronic version was found to have better structure.

Innovation and Author	Outcome	Evidence (even though often poor) or Commentary (no formal evaluation)
Introduced computer labs using Minitab (Stephenson, 1990)	Understanding Positive attitude	*Evidence*: The design involved a control (computer use) versus experimental group (no computer use) • Computer use was *not* found to increase statistical understanding as measured by increased exam performance when compared to a control group. • Perceptions as to the usefulness of the computer declined in the control group.
Flexible learning materials (electronic workbook) (Valacke, 1998)	• Enhanced understanding • Study postponement • Relevance • Exploration	*Evidence*: This is drawn from a number of studies. Study 1 • less postponement of study • success in course higher Study 2 No difference in study outcomes comparing face to face, electronic environment, a printed learning package and various options with an electronic workbook integrated with another medium
Interactive, multi-media modules (Yarbrough & Gilbert, 1999)	Bridging the gap between classroom and real world.	*Evidence*: • final grades were *not* influenced, but there were • improved attitudes
Strategy: Textbook		
Textbooks (Castro & Cobo, 1998)	Misconceptions	*Commentary*: The authors examined the treatment of correlation and regression in textbooks. It suggests misconceptions about correlation and regression, may be introduced by textbooks. It further suggests that writing a text book would require a deep epistemological analysis of the concepts." p. 675.
Textbooks. (Schacht, 1990)	Anxiety	*Commentary*: The author discusses the criteria for selecting a good text demonstrating the use of the Statistics Textbook rating Scale. Texts to reduce anxiety: 1. review basic algebraic operations 2. have a section on summation notation 3. include exercise answers 4. explain exercise answers 5. does not use definitional formulae 6. uses relevant examples 7. explicitly addresse student statistics and mathematics anxiety

Innovation and Author	**Outcome**	**Evidence (even though often poor) or Commentary (no formal evaluation)**
Textbooks (Sowey, 1998)	Aim to satisfy students learning needs	*Commentary*: The author discusses the choice of text and reasons for selection • short-term goals or long-term goals • gaining a higher grade; achieving a coherent understanding; being able to apply that understanding to real-world problems • retaining learning
Strategy: Themes		
Themes. (Cerrito, 1999)	Engage in statistical learning	*Commentary*: The author examines the types of issues that can be selected. Themes that have been taken for granted, *incontrovertible in students minds* are selected eg immunization. Students find, process and use technical information based on the tools of statistics to examine common beliefs.
Theme - risk (Dargahi-Noubary & Growney, 1998)	Motivation	*Evidence*: A comparison with a control yielded the observations •increased enrolments in follow on courses •attitudes and motivation improved •a greater proportion made conjectures and asked critical questions •more willing to participate and discuss •performed better on problem solving tasks •recognised connections between statistics and other interests
Theme - disciplinary (Shen, 1998)	• Broaden students thinking • Motivation	*Commentary:* The author suggests the use of interdisciplinary teams to provide different approaches to investigate a problem.
Theme - investment risk (Modern Portfolio Theory) (Shoesmith, 1998)	Motivation	*Commentary*: The author discusses ways of teaching through the risk theme.
Teaching statistical thinking using the Plan-Do-Check-Act learning cycle.(Wood & Wasimi, 1998)	Interest	*Evidence*: "Minute papers, end of semester written student assessments and anecdotal evidence have been used to monitor student reaction to the class. Our impression is that the method certainly captures the imagination and interest of more students" p. 171.
Strategies: Miscellaneous		
Identify learning styles (Bell, 1998)	Accomodate visual, auditory and tactile learning styles	*Evidence*: This is based on these groups performing differentially in a statistics subject.

Innovation and Author	Outcome	Evidence (even though often poor) or Commentary (no formal evaluation)
Learning arenas- problem based situations using the pose, collect, analyze, interpret and communicate cycle (PCAIC). (Ben-Zvi & Arcavi, 1998)	Meaningful learning	*Commentary*: The authors details the use of problem based learning areas to encourage meaningful learning in students.
Preparing posters in a cooperative group setup. (Denson, 1992)	Motivation	*Commentary*: Participation in contests was found to motivate students.
Co-operative learning activities to teach sampling. (Dietz, 1993)	Understanding and easier learning.	*Evidence*: This is provided by • student comments on course evaluations.
Self study workbook for teaching statistics to physiotherapy students. (Garven & Reed, 1993)	Knowledge scores	*Commentary*: Scores observed over four years provided a similar distribution to traditional teaching. They conclude that they could not determine if it lead to an acceptable level of understanding.
Co-operative learning (Gunawardena, 1998)	• learn without anxiety • improved performance	*Evidence*: No data but claims • that students are able to learn without anxiety • attendance improved • more willing to participate in class discussion • overall performance improved compared to the previous semester.
PACE approach - projects, activities, class lectures and exercises. (Lee, 1998a)	Emphasises: • active engagement in the learning process • relating statistics as a scientific tool for solving problems • teamwork • active involvement in report writing and oral presentation	*Evidence::* An opinion survey collected students perceptions of • perceived understanding of topics • the effectiveness of methods for analyzing data and report writing • degree of satisfaction with the activities used • adequacy of each activity The summary suggests students like the approach.
Teaching concepts in worked examples. Supplementary notes. Working in pairs. (Mahmud & Robertson, 1998)	Understanding	*Evidence*: This was obtained from observation, questionnaires, interviews and controlled comparative experiments. They found • 'Concepts [are] best presented within a concrete example but that the provision of notes is useful only when students work in pairs' (p. 736). Learning difficulties are associated with lack of prior statistics and too much material covered in the allocated time.

Innovation and Author	Outcome	Evidence (even though often poor) or Commentary (no formal evaluation)
Video (Moore, 1993)	Motivation	*Commentary*: The authors discuss how video can be used actively to promote discussion, ask about analysis and to introduce work. Moore's other writings are extensive and discusses other techniques eg using real world data sets.
Case study, stemming from a real marketing problem. (Russell, 1998)	• Reaction • Mean mark	*Evidence*: This suggested that when compared to traditional teaching the case study results in • more positive reactions and • better performance
Projects (activity based). (Ulmer, 1998)	Success rates	*Evidence:* • 'Median success rates for traditionally-taught sections is 57% as compared to 75% for the project/modeling section' p. 1265
(Searle, 1989)	Understanding	*Commentary*: The author focuses on the *why* and *when* not the *how* of statistics

4.3.3 Changes in Assessment

> Assessment is the process by which we gather evidence to help answer the question "How is it going" (Schloemer, 1997, p. 46).

It has been recognised that routine assessment might not assess many of the qualities considered desirable learning outcomes, for example the ability to generalise from the problems of the classroom to those of the real world (Garfield, 1995). Further it has been recognised by many educators that the form of assessment used can influence the approach taken by students to learning (Gordon, 1994), and indicate to students which learning outcomes are valued (Curcio & Artzt, 1996). Further, there are likely to be gender and ethnic differences in attitudes toward assessment (Forbes, 1998). It is also considered desirable for assessment itself to have educational value (Royer, Cisero, & Carlo, 1993) as part of the learning process rather than simply a means of evaluation (Huberty, 1998). Further, to effectively assess the impact of assessment or to make effective use of assessment practices one needs to align objectives, teaching and assessment. (Biggs, 1999; Crowley, 1997; Gal & Garfield, 1998; Garfield, 1995).

Garfield and Gal (1999) also provide a framework specifying issues to be examined when selecting assessment techniques. The assessor needs to identify:

• What to assess: concepts e.g. skills, applications, attitudes, and beliefs.

• The purpose of assessment: e.g. why the information is being gathered and how the information will be used.

• Who will do the assessment: e.g., the student, peers, or the teacher.

• The method to be used: e.g.quiz, report, group project, individual project, essay, or another method.

• The action to be taken and the feedback given to students: e.g. suggestions for revision and improvement, a summative grade, topics to review, etc. (Garfield and Gal, 1999, p. 7).

Three broad principles to guide assessment are provided:

The Content principle: Assessment should reflect the statistics that is most important for students to learn.
The Learning Principle: Assessment should enhance statistics learning and support good instructional practices.
The Equity Principle: Assessment should support every student's opportunity to learn important statistics.(Gal & Garfield, 1998, p. 773).

In further discussing these three principles Gal and Garfield (1998) propose eight instructional goals which need to be considered when determining assessment practices:

1) *Understand the purpose and logic of statistical investigations.* This goal is directed toward understanding the:

   ...big ideas that underlie data-based enquiries, and develop a sense for *why*, in what *circumstances*, and under what *assumptions* statistical investigations are initiated and conducted' (Gal and Garfield ,1998, p.774).

2) *Understand the process and phases of statistical investigation*. This includes an understanding of what is termed in Chapter 5, the research process. That is, understanding how to formulate a research question, planning the study and collecting the data, describing, exploring and analysing the data, interpreting the findings in the context of the research question and discussing the interpretations and implications.

3) *Master the procedural skills.*

4) *Understand the mathematical relationships*. This requires evaluating:

   ...if students recognize '*when* and *under what circumstances* certain statistical procedures have to be used, *why* they have to be used*, what* is their advantage over other procedures, *and what interpretations* can be attached to their results.(Gal and Garfield, 1998, pp. 775-776).

5) *Understand probability and chance*. Issues of concern include whether the formal understanding of rules of probability equate to understanding in other contexts.

6) *Develop interpretive skills and statistical literacy*. This may involve the combination of text with summary statistics or displays. It should also encompass whether or not students can apply their interpretive skills to other contexts such as interpreting the media.

7) *Develop the ability to communicate statistically.*

> Students should be able to effectively communicate, orally or in writing, about a range of experiences, technical issues, and results that come up during the investigative processes. (Gal and Garfield, 1998, p. 777).

8) *Develop useful statistical dispositions*. To be subsumed under the assessment of learning statistics this goal involves determining whether or not students *hold positive beliefs and attitudes* regarding their capabilities in statistical and probabilistic reasoning and interest in thinking statistically. Also examined is whether or not students *adopt a critical stance* or questioning attitude, and can appreciate the *power of statistics*.

They further elaborate seven assessment activities as suitable for achieving current goals in statistical education. These are:

1) individual or group projects;

2) portfolios of student work;

3) concept maps;

4) critiques of statistical ideas or issues in the news;

5) objective format questions to assess higher level thinking;

6) minute papers; and

7) performance assessment of statistical problem solving (Garfield and Gal, 1999).

The literature provides varied commentaries and some evidence regarding the use of these assessment techniques. For example, it has been found that collaborative group assignments using real data are enjoyed by students and are useful in creating understanding and interest in specific topics (Brunner,1997). Mapping techniques have been used in instructional planning, as a learning tool and for assessment purposes and teaching (Schau & Mattern, 1997). Students' concept maps have been used to reveal student understanding and/or misconceptions about statistics ideas allowing for this knowledge to be fed back into teaching (Roberts, 1996; Williams, 1998).

In this study the teacher researcher has used a variety of assessment practices. Much of the data forthcoming in Chapter 5 has been provided by a *reflective homework* wherein students were asked to reflect upon what they had done in class and the ideas generated and to write about it using an example of their own creation. The assessment techniques were directed toward students understanding and communication of the 'big ideas' namely variability as it impacts on the research process. As the teaching and curriculum developed increasing attention was applied to developing positive attitudes toward the use of statistics.

4.4 Educational Evaluation

In this section, a current and comprehensive framework is presented to provide the reader with a perspective on the evaluation of educational innovation. Unlike the positioning of the teacher-researcher, the framework was sought after the completion of the implementation phase and examined in the context of this work, retrospectively. The framework provided by Bain (1999) drew on a model for evaluating multimedia (Alexander & Hedberg, 1994). It provided a framework for the evaluation of 'all types of educational innovation.'

The Alexander & Hedberg (1994) framework includes evaluation during the design, development, implementation and institutionalisation phases of the project. The framework as presented in this section is to provide an indication of how and where in this study the criteria were addressed. The framework provided by Bain, postdates the collection of data. In accord with much of the grounded research methodology, the issues of the classroom, rather than models, theories or frameworks evident in the literature determined which data were collected.

4.4.1 Evaluation: Design Phase

Evaluation at the design phase has three components, curriculum analysis, teaching-for-learning analysis and the specification of the innovation. In Table 4.2 the focus of the evaluation and the purpose of the evaluation, both drawn from Bain (1999) are tabled, together with the sources of evidence and methods used when implementing the innovative statistics curriculum in Chapter 5. In this and other phases, the evidence provided does not necessarily encompass all that the frame-makers suggest. For example, other peers and experts did evaluate the effectiveness of the academic program. However the entire study was developed with a supervisor from an Education Faculty (Hedberg, from the Alexander and Hedberg, 1994 duo) and the professor-expert statistician. Furthermore the subject was developed, within a learning development centre and taught in an environment with daily discussions as to the progress of the subjects and students, and the strategies used to solve particular problems.

Table 4.2 Evaluation phase - design

Focus	**Purpose** **(Bain, 1999, p. 168)**	**Evidence and Methods**
Curriculum Analysis	'To describe the inadequacies in the current curriculum, with particular attention to the shortfall in student learning'.	• The work in Chapter 3 involved the sustained analysis of experts' thinking, including a mapping of a text/curriculum. • This analysis was undertaken in response to perceived deficiencies in statistical education as perceived by the lecturer, students and other teachers and as detailed in Chapter 1.
Teaching-for-learning analysis	To describe and justify the teaching/learning/ assessment process likely to bring about the desired learning outcome.	• The work in Chapter 4 focused on the development of conceptual understanding, an activity based/ experiential approach to learning or 'making learning possible' model was dominant. This replaced the traditional lecture approach based on a 'dissemination model' of teaching and learning. • A case study on assessment is provided in Chapter 5.
Specification of innovation	To describe and justify the proposed implementation, and indicate how it will facilitate the desired learning process and outcome.	• In Chapter 5 the actual implementation of teaching is detailed. The plan was to involve incorporate an activity based approach, eliciting student ideas, refining and developing the linkage of concepts and developing students' statistical language. • The aim was to have students 'see' or 'experience' statistical ideas and logic in action. Many other aspects of the teaching and learning environment are documented.

4.4.2 Evaluation Phase: Development

The development phase, differs from that in educational multimedia where students may be used to trial the multimedia products before formal use. In this study, evaluation during the development phase is essentially formative assessment, commencing at the onset of the implementation and concluding, temporarily, at the completion of the subject. Formative evaluation of each class was used to develop the subsequent curriculum and pedagogical techniques used in later classes. The summative assessment collected at the end of a subject became part of the formative process directing the curriculum and content of the latter implementations. Table 4.3 presents some of the evidence that was used at each stage of the evaluation in the development phase.

Table 4.3 Evaluation phase - development

Focus	**Purpose (Bain, 1999, pp. 168-169)**	**Evidence and Methods**
Formative monitoring of learning environment	To determine whether the innovation is functional in its context and accessible/attractive to students.	• Data were gathered and responded to throughout each implementation not just at the end of the subject. • Critical incidents that shaped the teaching process were documented. • Students were videotaped during a sample of lessons. • All work has been subject to informal commentary by colleagues within the program. One of the critical incidents later documented was in relation to a discussion with the Head of the program. • Both the lecturer and independent evaluators have collected student reactions to the program.
Formative monitoring of the learning process	To determine whether the innovation is influencing the learning process as intended.	• Within session assessment was frequent and formative in kind. Much of the evaluation in Chapter 5 is formative, examining the 'marking' of student assignments and the development of self-assessment. • Video clips reveal: students thinking aloud, student and teacher questioning and student reactions • A case study of the impact of assessment feedback on students has been examined. • A case study on the impact of assessment on the development of writing and attitudes toward learning has been included. • All work has been subject to commentary by colleagues within the teaching program. The subject has sustained its place in the program.

4.4.3 Evaluation Phase: Implementation

According to Bain (1999) the evidence sought in the implementation phase has two components:

- Summative evaluation of the learning outcomes. This involves 'evidence focused on the nature of the learning outcome, using outcome-relevant assessment tasks, supported by conventional assessments and student interviews where appropriate' (p. 169).

- Summative evaluation of the innovation's validity. Using evidence gathered in the development component and in the evaluation of learning outcomes, together with evidence on integration of the innovation into the curriculum, peer and expert review is used to judge the 'educational worth and viability of the innovation in the units/subject concerned (p. 169).'

The evidence relating to these two aspects of evaluation for this study is listed in Table 4.4.

Table 4.4 Evaluation Phase - Implementation

Focus	**Purpose (Bain, 1999, p. 169)**	**Evidence and Methods**
Summative evaluation of learning outcome	To determine whether the innovation is influencing the learning process as intended.	• A comparison of the assessment outcomes of three implementations of the subject was undertaken, using conventional assessment and concept mapping.
Summative evaluation of innovative validity	To determine whether the innovation is educationally appropriate in its immediate context.	• A case study examines, the origins of critical incidents- this subject, the overall program and everyday life - to determine their impact on the students' learning experiences. • Some evidence is provided on this subject effecting learning in other units (refer a case study on writing to learn).

4.4.4 Evaluation Phase: Institutionalisation

The evidence sought in this final evaluation phase involves two aspects:

- Impact evaluation. Evidence of the beneficial effects on:

> ...understanding and learning in related/subsequent areas of curriculum; indirect indicators (e.g. progress and retention rates); development of generic capabilities; transfer to the workplace (Bain, 1999, p. 169).

- Maintenance evaluation. This evidence involves:

> ...peer and expert review of the educational benefits of the innovations considered in relation to its maintenance and opportunity costs, and in relation to the educational and funding policies of the institution. (Bain, 1999, p.169).

The evidence that is available on these evaluation components is presented in Table 4.5.

Table 4.5 Evaluation Phase - Institutionalisation

Focus	Purpose	Evidence and Methods
Impact evaluation	To determine the robustness of the learning and its transfer beyond the immediate context of the innovation.	• The program was adapted and taught to different groups of students. A case study based on statistical literacy for law students looked at the extension of statistical ideas to new contexts.(Porter, 1998) • Evidence in the form of a case study, as to the improvement of generic skills, statistical thinking, and writing is included and published elsewhere.(Porter, 1998b)
Maintenance evaluation	To determine the sustainability of the innovation in the context of the whole course.	• The innovation was sustained for the remaining lifetime of the equity program in which it was embedded (5 years, 10 implementations). However the Gateway program did not survive University policy, restructuring and cost cutting initiatives resulted in the equity program within which the subject was taught being relocated to an external provider.

Many components of this framework draw upon evidence that has been provided through qualitative means. For example, various qualitative case studies were used to examine students' learning experiences and thus to provide a *summative evaluation of the learning outcome*. Similarly case studies were used to provide an evaluation of the *impact* of the program. Each of these forms of evidence is also subject to evaluation criteria as are deemed appropriate for qualitative studies. As qualitative methodology encompasses a wide variety of techniques and philosophical assumptions it is appropriate to examine the nature of the qualitative method used in the before examining the criteria for evaluating qualitative research.

4.5 Developing Grounded Theory through Qualitative Research

From the outset, the thesis has been conceived as a qualitative study rather than as a quantitative study. In the grounded theory approach *data are allowed to speak for* *themselves.* This may involve thick description of emerging data. Theory may emerge from through the process of categorising data, splitting and splicing categories and linking of data through methods firstly discussed by Glaser & Strauss (1967) but further elaborated by many researchers (for example, Creswell, 1998, Dey, 1993). It is through the working and reworking of data that the theory emerges in a manner that it is well-grounded in data.

In adopting a grounded theory approach, the participant-researcher's first step has been to position herself in relation to the theories and philosophies personally held. This has continued with the transition from unpacking statistical expertise to the second phase of developing, implementing and evaluating the statistical curriculum. In so doing, the researcher recognised that it is the researcher's orientation that causes him/her to adopt the project for study in the first instance. Further that the researcher's theoretical or philosophical viewpoints may direct the development and outcomes of studies, whether the researcher be engaged in qualitative or quantitative research.

The descriptive terms of the methodology that appear to best describe the process of unpacking statistical expertise was that of a participant-observer (statistician, psychologist and educator). Similarly the participant-researcher method (Anderson, 1994) provides an appropriate description of the approach to gathering data on the implementation of the curriculum, although here it may be regarded more specifically as that of a teacher-as-researcher paradigm. For it has been recognised that:

> ...teachers can also become their own recorders, composing and sharing their participant-articulated stories of practice. By then restorying their evolving perspectives, teachers may achieve more subtle and coherent understandings of classrooms and teacher self. (Diamond, 1997, p. 144).

Within the descriptions of developing grounded theory and teacher-as-researcher methodology fit many differing, but often similar, methods of collecting, analysing and reporting data. For example when discussing the use of reflective methods, Schon said:

> …[the] positions have several points in common. All of them purport to describe what is there, laying claim to a certain kind of objectivity. All of them adopt some form of ethnographic method, the careful qualitative description and analysis of case studies drawn from actual observation of an individual's or group's practice. And all of them represent

> their findings about practice in a distinctive way: They tell stories... [these] I shall call "mainfest stories of practice," "metastories of research on practice," "causal stories," and "underlying stories." (Schon, 1991, p. 344).

Many elements of different qualitative methodologies and data collection processes have been combined to develop what may be considered a grounded theory of statistical education. In the final phase of research, the early description that 'the researcher is the binding between the two components of the research' (Diary, 1995) was replaced by a recognition that the teacher-researcher's reflection plays a major role in making sense of the data. Initially interpreted as 'to think about', the issue of what it means to reflect arose during the implementation phase and is examined in the final chapter.

4.5.1 Teacher-as-Researcher Methodology

Paralleling the development of the constructivist view of the nature of knowledge has been the development of alternative research paradigms. Traditional educational research, conducted by academics and which separates theory from practice has been rejected by teachers as a means of improving their practice (Cookson, 1987; Doig, 1994; Gurney, 1989; Mitchell, 1985; Zeichner, 1995).

The need for an 'epistemology in educational research that regards inquiry by teachers as a distinctive and important knowing about teaching' (Cochran-Smith & Lytle, 1993) has been recognised for many years. One emergent paradigm being that of reflection-in-action provides a means of teachers conducting and determining the research issues to be explored. Similiar terms include that of the reflective practitioner, teacher produced knowledge (Zeichner, 1995), deliberation (Johnston, 1993), and as applied to education the teacher-as-researcher (Stenhouse, 1975).

The need for such reflective practice arises from what Schon called 'knowing-in-action', the sense that we possess knowledge of which we are unaware as 'our knowledge is enmeshed in our doing' (Smyth, 1986). It was this form of knowledge that the expert teacher was challenged to unravel in order to teach novices better (Bain, 1990). It is the teacher who is 'uniquely positioned to provide a truly insider or emic perspective on teaching that is not possible for others to gain' (Cochran-Smith & Lytle, 1993) and hence the suitability of the teacher-as-researcher for this thesis.

The teacher-as-researcher paradigm of research, recognises the value of teacher experience in the research process. Through a process of reflective practice, 'the practitioner as reseacher, moves from being in the midst of experience to exploring that experience, thus developing a kind of wisdom' (Adler, 1993). This is what teachers do on a daily basis to improve their teaching in the classroom. Teachers therefore are researchers:

> ...teaching *is* research and research *is* teaching. Teaching is more than simply the observable behavior; it is thought *and* action and the interaction of the two. The best teachers are researchers, able to systematically reflect on their own teaching. (Adler, 1993, p. 160).

This approach does not require, on the pretext of objectivity, the separation of the reseacher and practitioner, but rather maintains that the researcher should explore and understands the teaching experience.

> Rather than separate the knower and the known, research in teacher education can look to a knowledge base developed from practice...Pinar (1975) writes that to explore and understand educational experiences we must exist in them rather than removing ourselves from them. The charge to teacher educators, then is to reflect upon our experiences of practice and our inner worlds of meaning emerging out of that practice, in ways that are publically meaningful. (Adler, 1993, p. 160)

It can be viewed as an approach that contrasts with those which 'try and construct grand theories of the way the world works are no longer relevant' Smyth (1992, p11). In further elaborating:

> What we have in their place, are much more locally-based theories that recognise the idiosyncracies of site specific circumstances, and that acknowledge the integrity and worth of knowledge won by people at the workface....The emphasis in the reflective approach is upon practitioners being assisted to theorise their own accounts of practice, and how they might use that as a springboard for action. (Smyth, 1992, p. 11).

The issue is not 'what is best for practitioners to do', but rather 'what do practitioners need to know, and what do they already know or understand that might help them gain those extra insights?' Smyth (1992).

The methodology employed in this work includes not only that of the teacher-as-researcher but the aspect of action research that involves the teacher-researcher returning to participants with an interpretation of events so that they can provide further clarification and verification. Other formal and traditional modes of evaluation have also been included and used to support conclusions, about the subject's effectiveness, drawn by the teacher-researcher.

Evaluation criteria can and have been developed for teacher-as-researcher research, just as they have for other forms of qualitative research. For example, in choosing research to be read in a graduate program, the following criteria for evaluating teacher research were applied:

> ...the quality of the work itself (e.g. whether it is clearly presented, truly problematizes an issue, uses evidence to support conclusions), I am also looking for teacher knowledge which expresses particular points of view. (Zeichner, 1995)

4.6 Criteria for judging qualitative research

This current writing is to address the issue of whether or not the research undertaken may be judged as good research in the qualitative tradition. Traditionally, the separation of theory and practice has been viewed as the means of maintaining objectivity and hence valid research. However, as constructivist theory would argue:

> The general world of human experience, and the social and educational world in this particular case, is not an objective structure but a constructed, organic interaction of people organised and shaped by their culture, status and gender. Thoughts, feelings and action cannot be separated from each other, and traditional value free, non-ideological scientific objectivity is simply one fairly specialised form of knowledge, itself constrained under certain heavily constrained circumstance' (Jofili & Watts, 1995, p. 214).

In describing the writing of this thesis as a narrative or story, the researcher has intentionally alerted the reader to the proposition that both the framing of the study and the interpretation are based in subjectivity.

> *Story* keeps the story*teller* in the picture, suggesting that different narrators might make very different things out of the "same reality". Story connotes "just a story," thereby posing a question of objectivity (is there really anything out there independent of any particular observer's description of it?) (Schon, 1991, p.347).

Marshall suggested that the judgement of good research is necessarily subjective as different emphasis may be placed upon the different criteria by different researchers. From a constructivist perspective:

> Constructivist judges are happy with descriptions of the varied and multiple realities that are socially constructed. Good data are obtained by getting inside the worlds of others. The only truths are the emic realities of insiders. Certitude is not possible, there is no correct interpretation, and there is no end to the ever-involving interaction that reveals meaning. In the constructivist tradition, the debate over the goodness of one's interpretations is resolved by a dialogic process…The method is explicated in detail so that the reader can judge whether it was adequate and makes sense... The methods for attaining entry and managing role, data collection, recording, analysis, ethics, and exit are discussed. There is an auditability trail - Data collection and analysis procedures are public, not magical…There is abundant evidence from raw data to demonstrate the the connection between the presented findings and the real world, and the data are presented in readable, accessible form. (Marshall, 1990, pp. 192-3).

However, other researchers are not so inclined to accept the story-telling approach. Tooley and Darby (1998) write about the problems of qualitative method. They suggest that the key problem for qualitative research lay in its subjectivity, sampling bias and partisan nature. There is scathing commentary regarding the partisan nature of much qualitative work:

> The new research on teaching is highly political, in my view, and embedded with messages about power, justice, and inequality. Much of the information cannot be replicated because it is based on storytelling. Some of the data are suspect due to the limited number of subjects included and to the fact that there are too few checks for controlling a storyteller's social lens and politics. (Ornstein, 1995 p.125).

However, this leaves us with the challenge 'How do they [researchers] know what they claim to know?' and further 'What constitutes appropriate rigor?'(Schon, 1991, p. 343). From the earliest days in this study there was a struggle with these issues and an examination of the literature to ascertain what others considered as solutions.

There are numerous articles developing criteria for evaluating qualitative (Dey, 1993; Marshall, 1990; Schon, 1991). The practice of corroboration (Dey, 1993) or triangulation (Marshall, 1990; Tooley & Darby, 1998) are suggested as ways to establish the trustworthiness of the data. Triangulation may be described as follows:

> Triangulation is a way of cross validating research. It uses methods of comparison, to help assess the validity and reliability of the data collected. It can use several data sources or

> several data collection procedures, or a combination of these. If the data collected in this way disagree, then there is a dilemma about which the researcher is to believe - but this should then be a matter for making explicit that there is this disagreement in the data. (Tooley & Darby, 1998, p. 14).

In this study findings regarding students' understanding and their learning experiences have been triangulated using a variety of data collection methods. These included:

- traditional assessment of the statistics discipline;
- formal student evaluations;
- teacher reflection leading to course changes;
- learning exercises and/or questionnaire at the beginning of the session;
- questioning students about perceived learning difficulties mid session;
- questioning students regarding their interpretations of assessment feedback; and
- end of session learning exercises.

The study has also drawn on practices from action research methodology returning to the students, with teacher's and former students' interpretation of events, together with a request for confirmation or disconfirmation as they identify the challenges and experiences that they encountered throughout the course. Observation has been over a protracted period of time; the researcher's position has been presented in relation to assumptions and theories about how we learn. Recourse to the literature occurred throughout the study in order to find similar instances, and to identify how others had interpreted these instances. The recount of the study is as a story, with time and the process of unfolding an important component. It is supported by the sequence of diary entries, writings and conference papers presented. This provision of an audit trail allows the reader to address the plausibility of the story told about the research.

In the next chapter, the journey continues through the various implementations of the curriculum for cohorts of mature aged students preparing for university entry.

Chapter 5

Reflections In and On Practice

> Reflecting is at best an 'on-the-spot' action, a knowing response to an immediate situation but more often than not the knowing along with the moment of reflecting disappear. (Bamberger, 1991, p. 37).

5.1 Introduction

In this chapter, the journey continues. Here the focus is on *reflecting during* and *reflecting after* the practice of teaching. The immediacy of *reflection-in-action* varies from acknowledging critical incidents as they occur in class, the moments of reflection as new teaching is prepared, to thoughts about teaching and learning as the course ends and begins again. Reflection-on-action occurs some time after the action, removed from the teaching session. Throughout this chapter, that which is thought in close proximity to the act of teaching, when the teacher is emotionally intertwined with the activities of the classroom, the students, the materials and outcomes, is written in the first person. This may be in the form of a diarised indented quotation reflecting early, but *immediate* reflection (or the awareness of thought and action as it occurred in the classroom, but recorded immediately after). First person may also be used in the text referring to reflections occurring immediately or just after teaching. When reflection has taken place after the event, when names and faces have faded or when there is post-teaching re-analysis of the data, the subjective 'I' is replaced by the more distanced 'teacher-researcher'.

The data on which to develop a theory of teaching and learning statistics have been generated from observation of students, their activity, discussions, interactions, questions, responses, homework, examination results, subject evaluations and through documentation of teaching experiences by the teacher-as-researcher. The data have been collected several classroom settings. The first three successive implementations of a *Statistical and Mathematical Literacy* program for tertiary preparation students (*Gateway*) are the primary sources of data. For these case studies of the 'literacy' program, the unit of analysis is that of the classroom rather than that of the individual student.

The primary case studies are followed by a reflection on, and analysis of, specific aspects of these data regarding teaching and learning and on specific aspects of data collected in subsequent implementations of the Gateway program. These data have been used to examine:

- a case study of a student demonstrating the use of weekly homework in improving writing and statistical understanding;
- an examination of students' interpretations of feedback; and,
- an examination of the real world and broader *Gateway* contexts in which students' learn, and how these contexts impact on learning.

5.1.1 The three primary case studies

Foci of the three implementations

The development of the statistics curriculum for the 'literacy' program was based upon the process of unpacking statistical expertise in the teacher of the course and other experts, including the professor-expert as reported in an earlier chapter. These concepts were then restructured into a curriculum that had as its focus *variability*. During the second implementation, the statistical focus was extended and became *variation throughout the statistical/research process*. For the third implementation, no major changes were made to the statistical curriculum or exercises. However, the entire course was framed within the rubric of *learning how to learn*.

The primary case studies presented in later sections will convey how these thematic changes originated as well as providing a comparison of the learning outcomes from the three implementations. Table 5.1 provides a brief overview of the statistical units that were on offer in each implementation. The case studies will provide more detail on the content of the units and their manner of implementation.

Table 5.1: Statistical Units on Offer

Topic	Implementation 1	Implementation 2&3
You need no introduction to Statistics! Examining statistical ideas used by the general public	√	√
The Statistical Process: What questions would you want answered in order to accept a report as to environmental damage	X	√
Making Comparisons: Comparing seashells before and after an environmental event	X	√
Design: Testing the flight of planes	X	√
Measurement	√	√
Psychological measurement	√	reduced
Sampling: estimating mean age of the classroom population	√	√
Probability distributions: making decisions	√	√
Review of statistical concepts: concept mapping	√	√
Measuring Variability (orientation)	√ (Math)	√ (Stat)
Questions about Relationships	√	√

√ Units included; X units not included

As can be seen from Table 5.1 after the first implementation one unit was reduced (*Psychological Measurement*) and additional exercises (*Design, Statistical Process*, and *Making Comparisons*) were included to more fully expound the linkages between the statistical process and variability. The unit, *Measuring Variability* changed from a mathematical exercise on sigma notation to a statistical exercise on examining ways to measure variation.

The mathematics curriculum and pedagogy remained the same for all three implementations and is detailed in the case study of the first implementation.

Changes in teaching pedagogy

The teaching strategies used throughout all implementations have essentially remained the same. The strategies used will be detailed for the first implementation and exemplified by videotape for the second implementation. Additional time, three by $1\frac{1}{2}$ hour sessions, was made available for the second and third implementations. This allowed the inclusion of the additional statistics units. A constructivist model of teaching (the understanding of which will be detailed as they unravel) more explicitly drove the second and third implementations of the program. Associated with this model, more time was allocated to student reflection (through writing or group discussion) upon the class exercises in class time. Short video clips (10-12 minutes) selected from the *Decisions Through Data* series were introduced as a form of review and enrichment in the final weeks of the second implementation. Additional time throughout the term was also allocated to discussing the other learning difficulties students faced: dealing with uncertainty; anxiety; being perfectionists; feelings of being perceived as stupid; and difficulties at home when spouse or children had demands upon their time. In the third implementation a *learning to learn* framework was made more explicit to students. In the third implementation, the video clips were integrated from the second week of the subject and several learning exercises were introduced or modified.

5.1.2 Teaching in other contexts

Subsequent to the collection of the primary data, the statistical literacy component was modified and taught to different groups of students. One of these published elsewhere, *Statistical Literacy for Law Students* (Porter, 1998a) enabled verification of the type of understanding of statistics that can be drawn from this pedagogical and curriculum approach. Through a change in assessment techniques, it also allowed an examination of the issue of preparing students for life-long statistical learning.

The *Statistical and Mathematical Literacy* subject for pre-tertiary students and *Statistical Literacy for Arts Students* were each taught by another teacher on one occasion. Whilst not formally reviewed in this thesis, the second of these led to an exploration of alternative assessment mechanisms, group projects and peer marking.

5.1.3 Reflection on learning

At the time of writing this thesis, it was possible to re-examine data collected to follow-up the fleeting observations made in the midst of teaching and its aftermath. Questions were asked regarding the impact of feedback and the effectiveness of the feedback system, the nature of changes in attitudes toward learning mathematics and statistics, and the impact of the broader context of the tertiary preparation program and *everyday life* on the subject *Statistical and Mathematical Literacy*. Data that had been collected, including data on a fourth and fifth implementation of the tertiary preparation program, were available to address these questions.

5.2 Primary case studies, the participants: Students in the *Gateway* Program

The statistics program that is the focus of this thesis was developed for and taught to students in the University of Wollongong's tertiary entrance program called *The Gateway Program*. The completion rate for the Gateway, program ranged from 63 to 80 percent of students. On completion of the program, students were awarded a 'tertiary entrance score' and, with this, competed with other categories of applicants for entry to The University of Wollongong (NSW, Australia). Of the students who completed Gateway, approximately 90 per cent were offered places at the University, although not all chose to continue their studies.

The Gateway program was an equity program designed for men and women who have a history of being disadvantaged in their access to education. Priority was given to students who were classified as belonging to Department of Employment Education and Training (DEET) equity categories of students (rural or isolated, socio-economically disadvantaged, disability, Aboriginal and Torres Strait Islanders, women intending to study in non-traditional areas, and non-english speaking backgrounds). Many students fulfilled more than one of these criteria. Subject to there being places, a few non-equity students were also placed in the program. Of those who did not complete *Gateway* or later university studies, a high proportion was known to have done so for personal and medical reasons. Access to the equity program and University does not remove many of the difficulties, particularly medical and disability, associated with students' equity status on entry. From a mathematical perspective, the backgrounds ranged from that of students studying year nine mathematics twenty five years prior to Gateway entry to success in 3 Unit Mathematics at the HSC examination studied two years prior to entry in the Gateway.

The aims and objectives of the Gateway program in which the *Statistical and Mathematical Literacy* strand was taught were:

- to assist students in the attainment of the attributes of the ideal University of Wollongong graduate (University of Wollongong, 2000);
- to prepare students for entry to the University of Wollongong;

- to encourage students to accept responsibility for their own learning within a community of learners;
- to develop in students familiarity with acceptable standards of university work;
- to develop in students a realistic level of confidence. (Gateway program, 1994).

The Gateway Program comprised three strands:

- Language and Literacy including the subjects Writing and Speaking at University; Written Discourse; and Introduction to Literature;
- Statistical and Mathematical literacy; and,
- University learning including the subjects *Critical Thinking*, *History*, *Psychology* and, from the second implementation, *Intercultural Awareness* and *Library Skills*.

There was minor variation between implementations in terms of time, assessment and description of these subjects. Whilst these are seemingly disparate areas, there are, surprisingly, common themes throughout: learning how to communicate in different genres; learning how to comply with the demands of the system; and learning how to view knowledge from different perspectives.

The Gateway Program was 15 weeks in duration, with three hours of classroom contact on each of three days per week. One week was used for examinations and another for debriefing students, leaving 13 weeks for teaching. In the first implementation, eleven one and a half hour sessions were allocated to the mathematics and statistics strand. (This increased in later implementations). The contribution of this strand to the overall total Gateway mark typically varied between 25 -30 per cent of the total mark.

Each of the three implementations had approximately sixty to eighty students, divided into three relatively equally-sized groups. As with intakes of students into any university, there were variations in the calibre of the students and in this instance in the apparent disadvantage in terms of equity. The characteristics of the groups and the group dynamics were all likely to vary, with full-time workers dominating the evening group, and mothers with young children the morning and afternoon groups. Random assignment of the students was not possible and the classes were certainly not equivalent in many senses. The groups were based on student preference and availability of places. For this and other reasons, a positivistic experimental approach to evaluating the success of implementing the curriculum was not considered appropriate.

5.3 Case Study 1: Statistics as a study of variability

5.3.1 Introduction

In the primary case studies, statistics was taught within the *Statistical and Mathematical Literacy* program. The mathematics was essentially self-directed learning with testing and some demonstration in class. Mathematics was often woven into the undertaking of statistical exercises. This did not change throughout the three implementations. As the learning of statistics and the learning of mathematics were intertwined, it is necessary to examine the data in relation to both components, at least for the first implementation[1]. While satistics was embedded in the Statistical and Literacy program, the subject was embedded in a still broader framework, that if the Gateway program and *everyday life*. The impact of this is examined in a later section.

In this first case study there is:

- an introduction to the variety of pedagogical techniques used to teach both mathematics and statistics;
- an introduction to the mathematics and statistics content to be covered;
- a detailing of the various sources of data which were used to examine students experience of learning and their understanding of the statistical and mathematical literacy subject;
- an description of critical incidents in the classroom which shaped the teacher's perceptions of the teaching and learning experience
- student commentary on their learning experiences;
- teacher perceptions as to the nature of student understanding, drawn from the major assignment and subsequent concept mapping examination question;
- a reflection on issues which emerged during the teaching session, the use of language and experience in the classroom, the issuing of an invitation to learn, and metacognition;

---

[1]The use of the terms, first, second and third, implementation is used in the context of this study. The Gateway program had been in existence for several years prior to this implementation. However, coinciding with the commencement of this study was a major revision of the Gateway program, including the addition of a statistical strand to the existing mathematics component, and my commencement of teaching within it. Since commencing, I taught this subject in Gateway on nine occasions, with another teacher teaching it on one occasion.

- the nature of feedback from colleagues upon the pedagogical approaches implemented; and,
- problems which were suggestive of a need to change aspects of further implementations.

5.3.2 The Mathematics Component

In working with this group of students for the first time, what was assumed or known about them was that they came from diverse mathematics backgrounds, many would find mathematics difficult, some would find mathematics easy and a few would have prior formal knowledge of statistics. It was assumed that the cohort of students would be similar to that of the previous year (taught by another lecturer). This cohort would be expected to show competency in targeted mathematical skills of a standard comparable to Years 7 and 8 in NSW high schools. The desired aim for the mathematics curriculum was to produce:

- students who were responsible for their own learning.
- students who had sufficient mathematical skills to undertake a 100 level social science or humanities course; and more specifically.
- students who were capable of performing addition, subtraction, division, and multiplication with whole numbers, directed numbers, decimals, fractions and indices. They were to know the conventions for ordering mathematical operations (for example, the use of bracketing to denote priority of calculation) and to understand elementary mathematical notation and formula substitution. In addition, they were to correctly use sigma notation.

The meagre time allocation in some sense 'forced' the approach to teaching and learning. There were nine teaching sessions, plus one for an examination and one session for debriefing. Approximately one hour per session was allocated to the statistics strand with the remaining half an hour for mathematics and other matters as they arose in class. There was, however, no distinct break between the learning of mathematics and statistics as the teaching and learning of mathematics could often take place in the context of the statistical lessons. It was reasoned that the available teaching time would be insufficient to teach all the mathematics required by 100 level subjects. Furthermore, students seeking entrance to mathematics and engineering studies would need to complete a 2 Unit mathematics preparatory strand elsewhere. Survival in the university culture involves more than knowing content and skills. It also involves:

- finding and using resources;
- knowing how to undertake tests (both open book and closed);
- managing stress;
- submitting work of appropriate quality on time;
- being open to doing what one finds difficult or what one does not want to do;
- showing up to class; and,
- analysing questions and content to determine the most important concepts.

It was reasoned that students could, at best, be set upon the path of learning their mathematics in such a manner that they would feel confident that they could continue to develop skills. It was decided to implement a self-directed approach to learning mathematics, - albeit a closely monitored system.

The approach taken to teaching mathematics was premised on there being two different groupings of students in the program:

- students who felt mathematically able. Whilst not recently using much of their prior formal mathematics, these students would be confident, needing only to refresh their skills. One of the anticipated difficulties for this group was the possibility that the mathematics, being too easy, might induce boredom.
- students who had poor experiences in mathematics learning. It was reasoned that for these students, a traditional classroom teaching style, primarily with demonstrations on the blackboard and lots of worked examples, would probably reinforce that previous poor experience and that an alternative should be found.

Data collected from students during the first meeting and at the end of the course confirmed the premise that there were these two types of students and that there was a diverse range of mathematical skills. Neither the details of students' mathematical histories nor my teaching experience at university level prepared me for students who could not perform subtraction of whole numbers or for students who recollected when multiplying by tens or hundreds that 'in multiplication one did something with a nought.'

It was not a surprise to find that many students saw no reason as to why they should undertake a mathematics preparatory strand. These students were aiming to undertake courses in subjects other than mathematics or statistics, subjects such as psychology, history, management or english literature. However by way of introduction students were made aware that, at this university, all graduates were expected to become statistically literate, and furthermore that subjects such as psychology or nursing did, indeed, involve mathematics and statistics.

The *teaching of mathematics* involved establishing a structure within which students were to work and learn, and through which they would be able to communicate and receive support for their mathematics learning. Elements of this structure included:

- resources provided for students;
- support available to students;
- a system of self-assessment;
- demonstrations in class;
- mathematics homework;
- mathematics practice within the statistics program; and,
- a major mathematics assignment.

Resources. The primary resource was a set of mathematics booklets covering all topics, (with the exception of sigma notation). These were self-marked (although some marking was reviewed by the teacher) and submitted when completed. There were no marks attached to the completion of this work. Students commented in the workbooks if they experienced problems or if they wanted additional resources. They could also indicate that they wanted a tutorial session with the teacher. In response, the teacher explained the solutions and/or provided an appropriate alternative handout on topics found difficult. The teacher was able to monitor the work students were undertaking through examining the submissions.

Support. Students could ask for help outside class time, although they were encouraged, firstly, to explore an alternative text and other resources.

Mathematics homework. During the first week, students were asked to find a textbook 'at about the Year 7 level' so that they could work through exercises on the types of questions they had got incorrect or for which they had little confidence in their methods of completion. These questions were submitted for comment and solution if the student had difficulties. Those who had no difficulties in the tests were not required to submit practice examples.

Demonstrations in class. After the fifth test, time previously allocated to 'testing' was used to demonstrate different examples. This took approximately 15-20 minutes each week. At this stage most students were actively engaged in the process of directing their own mathematics learning. Whilst some still dreaded the test, others indicated that they waited for it in order to test their skills.

Mathematics in statistics. A number of statistical exercises involved mathematics, providing the opportunity for further skills practice.

A system of self assessment. Commencing in week one, students were given the first of three *closed-book* tests. After this rigorous start, two open book tests followed, allowing students to use what I term a *manual-based* approach to learning. Provided students recognised the problem, they could utilise notes they had made to solve problems. After each test there would be a brief period working through solutions. *Teacher talk* stressed that evidence of learning should be seen as change from where the student commenced. The tests gave students the opportunity to practice their examination technique. The tests generally incorporated a mix of questions drawn from several topics. The tests were used for self-diagnosis and for the teacher to monitor the level of assistance required. Many students appreciated the simple level of diagnostic summary of the tests as demonstrated in Table 5.2. A cross (x) was used to indicate that use of the skill was incorrect. A tick (√) indicated correct use of the skill. A given question could involve more than one skill. The pattern of xx√√√ could be used to identify where students had learned what was previously unknown. Marks were attributed, not on the outcome of the initial tests, but the skills in evidence in the later weeks, that is, only for the sequences of skills ending in ticks. Students were rewarded only for the skills evident at the end of the subject. For the skills shown in table 5.2, the student would have lost marks for the sequences ending in a x, that is for the skills addition, subtraction and multiplication of decimals (this was an unusual scenario).

Table 5.2 An abbreviated skill development sheet

skills	test 1	test 2	test 3	test 4	test 5
meaning of numbers	√√	x	√	√	

rounding	√x	x	√		
addition & subtraction	√x	√	√		
addition decimals	√	x			
subtraction decimals	√	x			
multiplication	√				
multiplication of decimals	√	xx			
division	√	√	x	√	
division by decimals	x√√	√	x	√	
division in decimal form		x		√	
adding & subtracting fractions		xx	√	√	
multiplying fractions		x	x	√	
dividing by fractions		x	√	√	
sigma notation		√√	x	√√	
order of operations		xxx	√√√	√√√	
fractions to decimals		xx		√	
decimals to percentages		x		√	
substitution in formula					
TOTAL	10/14	4/22	8/12	14/14	

Mathematics assignment. The major mathematics assignment required students to draw up an examination paper, provide the solutions and to make notes on how to answer the questions. The rationale for this assignment was that students could gain experience in preparing for an open book examination and the assignment could be used in the final examination. Once again, many students appreciated the form of this assignment, and did it conscientiously. The assignment had been limited to 15 questions: however, students mostly wrote several parts to their 15 questions! For the most part the difficulty level of these tests was appropriately set, although coverage generally needed to be more encompassing as students often omitted major topics. Written feedback on the assignments identified for the students the areas that had been overlooked and the areas where they had presented examples which were considered too easy, too difficult or incorrect.

Video segments on teaching mathematics were prepared during the second case study. However, this footage characterises the teaching of mathematics in all three primary implementations. As the viewing of the clips (refer the CDROM attached) requires a move to technology, all clips have been collected and discussed in Section 5.4.4.

5.3.3 Data Collection: Mathematics and Statistics Learning

Data regarding the learning experiences of students originated from:

- the teacher's observations and reflections, generally recorded in some form of diary or formal writing, on what transpired in the classroom; and,

- the students' written responses to questionnaires.

Teacher Generated Data: Observing the impact of teaching on students

The foci of the teacher observations and reflections included:

- student discussion;
- the asking of questions;
- the answering of questions;
- various verbal and non-verbal cues that originate in the interactions in the classroom; and,
- critical incidents in the classroom, incidents that had some impact upon the teacher, either reaffirming the teaching stance or inducing change.

These were all documented. Other data that allowed examination of the changes throughout the subject have been generated by students through questionnaires collected by an independent unit or through student assessment.

Student Generated Data: Learning Experiences

At the start of the teaching session, students were asked to complete a questionnaire. This was to provide the teacher with information as to the students' prior mathematical background, students' awareness of preferred learning approaches and their level of confidence as they anticipated undertaking the subject, *Statistical and Mathematical Literacy*. The questions were:

- Q1. So that I have a better understanding of the needs of students in this class I would like you to describe your mathematical background. What year and level mathematics you have completed, how long ago and where was it completed? How do you feel about yourself doing mathematics?
- Q2. Think about the situations in which you learn best and those situations where you learn worst. Tell me, if you can, those sorts of things that you think will help you better learn statistics and mathematics, and those things that will hinder your learning.
- Q3. On the following scales please circle the appropriate response for you.

  (a) At the thought of having to do mathematics I feel:

Not at all confident								Highly confident	
0	1	2	3	4	5	6	7	8	9

(b) At the thought of having to do statistics I feel:

Not at all confident								Highly confident	
0	1	2	3	4	5	6	7	8	9

- Q4. Is there anything else you would like to tell me before we commence this course?

At the end of the session, in the debriefing week, students completed two evaluations. The first of these was collected by the teacher and was directed at ascertaining the strengths and weaknesses of the course. Whilst students were not required to provide their names, most did, allowing a matching of data with that collected in the first week and throughout the session. The teacher asked students to respond to the following prompts:

- Take a few minutes to think about what you have learned in the Gateway program and particularly the subject *Statistical and Mathematical Literacy*. What are the strengths and weaknesses of the overall program and the Mathematics / Statistics program?
- How would you improve the Mathematics / Statistics section?
- What have the landmarks been for you during this course? When did you take a step forward or backward in your learning in the course, and more particularly, in Mathematics and Statistics?
- Based on your performance in this and the other strands of Gateway do you feel ready at this stage for university studies? In what areas do you need further preparation?

Students were again asked to rate their level of confidence in learning mathematics and statistics.

The second evaluation carried out by an independent unit was part of the University's formal assessment of courses and lecturers. The lecturer was provided with a summary at the end of the teaching session when student results had been finalised. Students completed this anonymously. Further, it was not possible to match student responses on the various items. Students were asked to rate their experiences on the following scales:

The lecturer sequences material in the unit, very illogically (1) to very logically (5)
The quantity of material has been, far too little (1) to far too much (5) (3 is the preferred response here)
The lecturer stimulates me to think about the subject, never (1) to always (5)
The lecturer has demonstrated understanding this subject, very poorly (1) to very well(5)
I have felt enthusiastic about attending lectures in this subject, never (1) to always (5)
The subject material has usually been, very uninteresting (1) to very interesting (5)
The lecturer has presented the material, very unclearly (1) to very clearly (5)

This lecturer's interest in assisting me to learn is, non existent (1) to very high(5)

The data for these items allowed a quantitative comparison of students' reactions to the course over the three primary implementations. This comparison is provided at the end of the third implementation in Section 5.5.

5.3.4 Data Collection: Mathematics Understanding

The formal data on students' performance and understanding in mathematics was derived from *resources* detailed earlier when discussing the structured learning environment for mathematics. These included:

- closed and open book mathematics skills tests;
- mathematics workbooks and additional exercises as the student felt necessary;
- the major mathematics assignment, where students devised and completed their own examination paper; and
- a formal open book examination in mathematics.

These provided a profile of work for each student that could be examined from the perspective of the understanding and learning of mathematics.

5.3.5 Mathematics outcomes

Critical incident analysis

As the teacher-deliverer, at times I felt uncertain with this approach to teaching mathematics and only just kept myself from taking back the responsibility of learning from the students. Small things, comments, and the encouragement of both my fellow staff members and students kept me persisting with the teaching method. Some of the landmarks over the eleven weeks of mathematics are presented in Exhibit 5.1.

Exhibit 5.1: Critical incidents in 'teaching' mathematics

Week	Critical Incidents in 'teaching' mathematics
Week 1	
Activity:	Maths pre-test followed by demonstration of solutions.
Students:	Swapped papers and marked them.
Teacher:	Over 50 per cent the students had difficulty with long division, 30 per cent with multiplication, 15 per cent with subtraction, 10 per cent with addition.
Week 2	
Activity:	Statistics exercise (sampling) involving maths (addition and division).
Students:	A few students arrived having found and used resource books.
Teacher:	Handouts provided on multiplication and long division.
Week 3	
Activity:	A longer and more difficult test. Solutions demonstrated.
Students:	One could feel the tension, hear the expletives under students' breaths. Appalling test results.
Teacher:	The test was stopped midstream & students asked to take note of their physical state. They were reminded that this test was for them - so that they could see any improvement in their skills & gain experience in taking tests. Calm returned. To reinforce this theme, students were then asked to mark their own class tests prior to submission. This practice continued later in the subject. Many were still not working through the examples in the mathematics workbooks. Both the students and teacher were questioning: 'Will this approach work?'
Students:	One student told me with great delight that the questions she had practised she got correct and the ones she had not worked on she got wrong.
Teacher:	I felt that this student had experienced the power that comes from teaching oneself. I refrained from taking the responsibility for learning back from the students.
Week 4	
Activity:	Test. Demonstration prior to the test. Solutions left on board.
Students:	The word 'yipee' written on a test paper as another found success.
Teacher:	Many students prepared for the rigour of closed book tests; the process of self-direction started. Some students still not working appropriately.
Students:	Four students identified themselves as showing little improvement. A few were concerned with the need to memorise formulae and procedures. (One of these four later came in the top group of students.)
Teacher:	Future tests were open book, in order to allow students to develop and show skill through using resources such as prepared notes and manuals.
Activity:	Statistics exercise in measurement and in developing a psychological test, to examine learning.
Teacher:	This exercise involved developing a test to measure likelihood of success in mathematics in relation to work habits. The exercise opened up discussion as to how students were approaching their work. Were they working through examples, summarising theory, doing practice spaced over time or all in one sitting etc.
Week 5	
Activity:	Open book test after a two week break. Reflection on activity.
Teacher:	A divergence in student results: most students had completed student workbooks; others had not.
Students:	Another student was identified as being in difficulties. This student wanted all maths to be taught in class. As a full-time worker she found there was not time to do the extra work.
Teacher:	I perceived that a student was angry. No practice examples had been submitted by this student.

Week	Critical Incidents in 'teaching' mathematics
Students:	Other students offered support for my approach to the teaching of mathematics.
Week 6	
Activity:	Demonstrations; then an open book test.
Teacher:	Perceived that the test was delivered to two groups - those who were now looking forward to the test as a measure of their growing skills, and those for whom another approach was necessary. Mini tests had served their purpose.
Week 7	
Students:	Student made a telephone call to ask if he could use a calculator as he had been allowed to do so in other courses. He was proficient at using a calculator.
Teacher:	This student projected all sorts of reasons not to learn his math - too much work, he was too slow, he would forget, he worked full-time and did not have time. The response to his request to use a calculator was 'No, I am not just interested in your maths. One of the things that I am interested in is, how you learn to do something that is, for you, difficult to do'.
Week 8	
Activity:	Demonstrations.
Students:	The student who requested a calculator arrived and proudly showed his manual for learning maths.
Week 9	
Activity:	Last demonstrations before final exam.
Students:	Student arrived for individual tuition as she had requested. No longer angry, resource books obtained, practice started, individual help no longer requested.
Teacher:	Head had supported my stance that prospective university students should be able to work through a year 7 textbook. Another lecturer delivered a lecture on coping with confusion. The approaches, of blaming oneself and blaming the teacher, for the confusion were examined.

The old hands who teach the *Gateway* students talk of ‘the students coming good. For some this is early, for others in the last week’. So it was in this course. Apart from the content of mathematics, students had to learn to comply with course requirements. Many students initially found it difficult to attend to homework. The first homework had a compliance rate of less than 50 percent. The major assignment was expected to be on time. Compliance with this was reasonable, with three non-submissions and four late out of the then total of 51 students (47 students finally completed the subject). One student commented on the difficulties associated with assignment deadlines as follows:

> My problems are not so much... the subject matter in Gateway, but rather on trying to modify my lifestyle as a whole to include things like deadlines, penalties, arguing as devil’s advocate etc.

Subjecting students to the anxiety of tests, requiring them to direct their own learning, dealing with the attendant anger could have been a story of disaster. When asked about the strengths and weaknesses of the subject, many students commented on the weekly tests and homework. Four such comments were:

> The weekly assignments and regular revision tests in maths and stats gave me the feedback I needed so I could improve in areas that I was not good at.
>
> Strengths - homework and tests (weekly) that were marked and commented on were very helpful for getting through the program and the Maths/Stats program.
>
> Strengths: Booklets, skills tests, explanations and showing on the board.
>
> Strengths: Skills tests - revealed that I was understanding what was being taught. Step forward every lesson.

Yet another student commented on the importance of being able to mark her own paper, as she then felt she could attempt the questions and not be embarrassed about her silly mistakes.

As anticipated some students found the mathematics too easy. The students with the most mathematics potential, in terms of their mathematics background, were not extended in their mathematics (although they were by the statistics) as exemplified by the following suggestion from two such students:

> After the initial skills test divide the group into say two groups. Let the advanced students go for it (with a supervision schedule). The teacher can then concentrate on the weaker students.
>
> I feel that if there were more emphasis placed on the statistical aspect of the course (especially r scores and t-tests etc) that this would be more beneficial than 6th grade mathematics, that could be learned by students at their leisure… I went forwards when I realised I knew more about statistics than I thought I knew; and backwards, when in psychology, the results of our experiment were returned to us, including a t-test result. This aspect of statistics has totally confused me and I am still none the wiser.

Some students clung to a vision or expectation of what the course might be. They confronted the actual course only when they let go of this expectation. It was at this point that their work began to improve. Those who found the path early succeeded. For those who found the path just a little later, outcomes depended upon the extent to which their mathematics background was impoverished. For example, in the listing of critical incidents the teacher was aware of two students whom she perceived as resisting the system of mathematics learning:

- The student who requested a calculator (as identified in the record of critical incidents) had started the session with the comments, 'mathematics makes me nervous', and, 'don't yell at me if I get it wrong'. For this student the course was 'too fast'. At the end of session the student said, 'I have learned that running at maximum stress and maximum work load does nothing to promote an appreciation for maths. The speed that maths has been presented has only caused confusion.' He finally concluded that he needed 'a course in stress management'. This student resisted the system, wanting to use a calculator[2], and hoping to gain permission to do so. Commencing work in the eighth week left too little time to master the mathematics, although he did ultimately take responsibility for working through the workbooks. With few initial mathematical skills he did not have time to master the skills covered in the subject.

- Another student indicated at the beginning of session that she enjoyed mathematics, although she had completed schooling some thirty seven years ago and at that stage had not learned algebra. She was identified in the critical incidents in Week 5, as the angry student. Her earlier comments indicated that she would need to be 'shown examples'. The teacher felt she had, at last, started her mathematics work in the ninth week and interpreted this as the student 'no longer resisting the system of working through the workbooks'. The student had, in fact, 'finally employed a tutor and learnt a lot in a short time.' With reasonable mathematical skills she was able to do this successfully but she remained discontented with the teaching, indicating that there was 'insufficient teaching'.

For those students who were not so mathematically able at the commencement of the course, the learning experience varied depending upon whether or not they could adapt to taking responsibility for their own learning and do so early in the course. Another small group of students responded to the challenge by seeking out others, such as friends, fellow students or some one-on-one with the teacher to teach them, as they had been taught in school. However, they did not cling to a belief about what should happen in the classroom, preventing them from starting their workbooks. These students worked some of the materials themselves, discovered what they could master before seeking help, unlike the students identified in the critical incident schedule.

For many students from poor mathematics backgrounds the experience of learning mathematics in this manner was overwhelmingly positive. Comments from three of these students:

---

[2]A choice was made that students should learn how to perform mathematics skills without the use of a calculator. The calculator could be used performing the tedious mathematics as it occurred in the statistics section and to check their work. At what stage calculators should be introduced into the teaching of mathematics is subject to considerable debate. There would be benefits from teaching students how to use calculators before they moving into disciplines which utilise statistics. However, the building of confidence in students' own abilities to demonstrate mathematics skills was adopted at this stage. Many of these students needed to acquire skills, such as being able to estimate the size of expected answers before utilising a calculator.

> I recognise my maths ability as basic skills level, so I found the course terrific, increased my confidence and ability. Surprisingly, for the first time in my life I actually found maths interesting and challenging. Thank you.
>
> This stats and maths subject was the first time I have ever understood maths and became even to enjoy it.
>
> The whole *Gateway* course was a great leap forward. It taught us all to slowly open our minds and gave us the confidence to believe in ourselves. In particular maths and stats I found to be paced, easy to understand and always at a level with which I could cope. Never having liked or understood or remembered a single mathematical concept before, I found myself enjoying *Gateway* stats and felt a degree of confidence as I sat my first maths exam.

These and other students gained by finding strength within themselves to learn what was for them difficult:

> The strengths were that if you really wanted to achieve you had to do it on your own. I found, for me it was much better to work independently rather than to depend on others. My greatest weakness is confidence. I feel more confident now, to know if I look up my information, that I will learn independently on my own.
>
> I know that I still have a long way to go but I have learned (finally) how to approach maths/stats. It has always been a subject that I have struggled with and with more effort I know I will improve. Maths was something that I was never comfortable with, but to get a pass is a great achievement for me. I don’t think you can improve the course. I found the maths/stats much easier to learn at *Gateway* than I did at school.

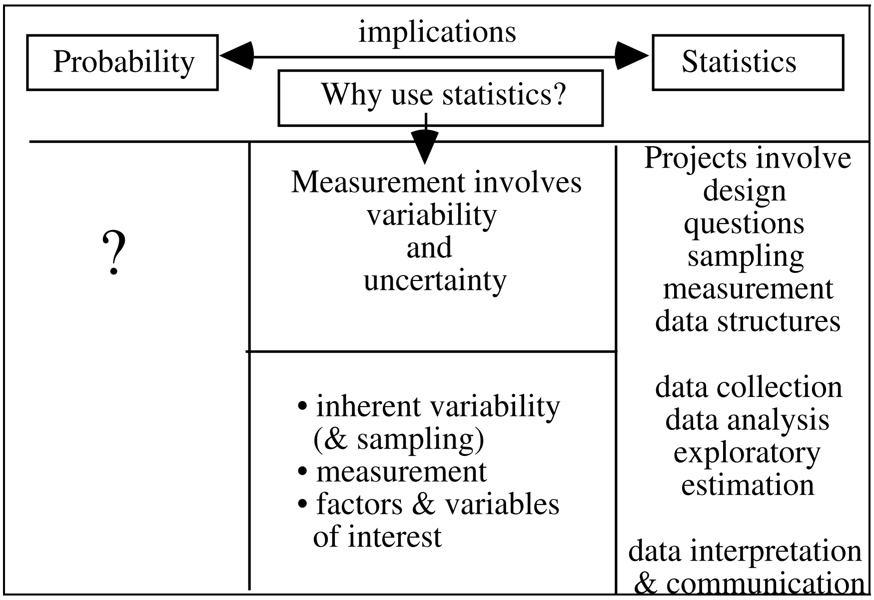
What was encouraging was that in the students evaluations the majority of students indicated 'in this subject (mathematics) my enthusiasm has increased or increased greatly' (n= 27 students, 67.5%), 'remained the same' for 10 students (25%) and 'decreased' for only three (7.5%). There were seven students either absent from the evaluation or for whom it was not possible to match the pre- and post-questionnaires. In less than half the time allocated to teaching mathematics in the previous year, the students covered five of the seven workbooks and, in addition, completed a major statistical strand. The mathematics level that most had attained should hold them in good stead for undertaking 100 level social science, education and humanities courses. Whilst they did not have all the mathematical skills necessary, most students had taken responsibility for their own learning in mathematics and in so doing they had acquired a confidence that they could continue to work through the skills they did not know.

5.3.6 The Statistics Component

Statistics Curriculum

For this implementation, the theme was *statistics is a study of variation* in the world around us. However, the teacher-researcher's experience and the work in unpacking expertise suggested variability was rarely portrayed as the central concept in teaching statistics. Unpacking to date had created a representation of major statistical ideas as portrayed in Figure 1. The role of probability had not been fully elaborated, just as it is often omitted from statistics curricula. The teacher-researcher perceived statistics texts and the courses on offer to non-statistics majors (such as those in commerce or psychology), as concentrating on the statistics side of this diagram (Figure 5.1) with statistics becoming a sequence of tests such as t, chisquared and ANOVA.

Figure 5.1 Positioning variability and probability in relation to the teaching of statistics.



Teaching units were prepared to address the 'why' of statistics: that is why we are concerned with variation in sampling, measurement, description and analysis and decision-making. The statistics teaching weeks were sequentially arranged as in Table 5.3.

Table 5.3 Sequence of statistics teaching

Week
1. Mini lecture: An introduction to Statistics (Exhibit 5.1)
2. Exercise: Sampling (Exhibit 5.2)
3. Exercise: Physical measurement
4. Exercise: Psychological measurement
5. Mini lecture: Questions of relationship
6. Review: Providing a conceptual framework
7. Exercise: Probability distributions
8. Review
9. Review
10. Exam
11. Debrief: Linking statistics to the next curriculum

The focus in the statistics component was to be on concepts rather than mathematical statistics and formula, that is calculations. The measurement of variation was taught through the mathematics program.

Teaching Methods: Statistics.

The course used a mix of lectures, activity-based exercises, and review or ‘getting it together sessions’ as follows.

Lectures. Students were given two mini-lectures. With these classes there was an initial statement of expectations: ‘if you do not talk to me, I will not talk to you’. The lectures were based on providing scenarios and asking questions about them. (Refer Exhibit 5.2). At times, when the responses were too few, or too soft, deafness was called into play. The questions were not rhetorical questions; students needed to provide answers to them and they did so. The first of these lectures seduced students into responding with ideas and sometimes terms that were statistical and by so doing some began to accept themselves as statistical thinkers. The lecture was designed so as to initiate a connection between statistics and everyday life. For some, the connection was far more difficult to establish. The lectures are essentially carefully crafted dialogues between lecturer and students.

Exhibit 5.2 Mini lecture - You need no introduction to statistics!

Classroom scripts: You need no introduction to statistics!

Teacher's script: We all use basic statistical processes in our everyday life. Students' response: Disbelief!

Teacher's role: Recognise the students' expressions; reflect and refute students' expressions.

Teacher's script: Little Johnny comes home from school with a 2/10 and you begin to worry....Why?

Teacher's role: The students will provide the answers if challenged and allowed time to respond. The answers generally relate to the anticipated distribution of marks and particularly the anticipated mean and range of marks.

Teacher's script: Why ? What else might you have thought?

Teacher's role: There are no wrong answers - just answers collected at the wrong time. Encapsulate their ideas and label them with the appropriate terminology - eg. mean, median and range. Let them know you will pick up some of the unused ideas later. The key ideas are generally produced. If not, story-telling comes into play as follows, 'Last time I taught this lesson one of the students…'

Teacher's script: We assume that the class average *(the mean)* is 5 or maybe even higher, or the score that most children received was 5 *(the modal score)* or that the *range* of scores was from 1 to 10. We use ideas, such as average, mode and middle score *(median)***.** We use ideas such as how far the marks are spread out - *the range.* If we find that the class marks in Little Johnny’s class *range* from 1 to 4 out of 10 with the most common score being 2 - are we still worried about little Johnny? Why not? (Students may have other concerns about the teaching).

Students' elicited script: No, because the assumptions were incorrect.

Teacher's role: Encapsulate the ideas, echo them and use appropriate statistical language.

Teacher's script: Our *estimate* of how well Johnny was doing was not based on the *correct assumptions* about the *distribution* of the class marks. Marks did not range from zero to ten and the typical mark was two marks, not five. The decision to worry was based on a set of incorrect assumptions. (And of course there is still the worry about a teacher whose marks average two out of ten.)

Teacher's role: Still disbelief that they think statistically. Reflect their disbelief back to them. 'You don't believe me yet, but lets go on'.

Teacher's script: These are statistical ideas. Whilst we use many statistical ideas we don't often recognise them dressed in statistical jargon. Possibly we won't always know the formula or how to use it to come up with the mathematical numbers or patterns or statistics that represent these ideas. But let me go on… Is there anyone here who has worked on the school P&C? or as a fund-raiser? Recently the primary school, my children attended had a centenary celebration and they organised to sell specially designed T shirts. How would you work out how many shirts to order for each size (6,8,10,12,14)?

Students' responses: Students provide a variety of responses including conducting a *sample* and *census* of students and parents.

Teacher's script: Suggest they ask the children and play act if need be? Why not ask the children? Ah, I'll have three Miss - a red one, blue one and a black one!!! Perhaps you would send home a note to find out how many shirts each family would buy, and what size they wanted. And then what would you do when the notes came back from home?

Classroom scripts: You need no introduction to statistics!

Students' response: Students generally indicate that they would keep a tally.
Teacher's role: Role-play getting the orders from different students so as to develop the tally chart. Then show them the statistician prepared tally chart. It is the same as their chart.

Overhead 1: A tally chart of shirts to order

Size 6	11111 11111
Size 8	11111 11111 11111 11111
Size 10	11111 11111
Size 12	11111 11111 11111 11111 11111 11111
Size 14	11111 11111
Size 16	11111 11111 11111 11111

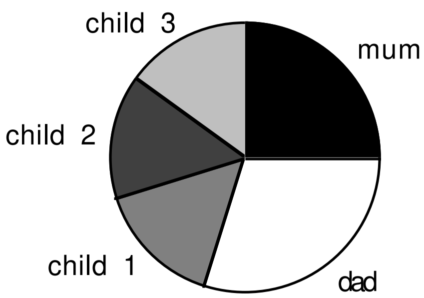
Teacher's script: So in other words you drew up a *tally chart.* Is this how you would present it to the principal? Okay, so for each size we add up the tally marks. So perhaps you would provide the principal with a *frequency distribution* to represent the number of each size of shirt that you wished to order. Overhead 2 (omitted) is as the students expect, but it is labelled as a *frequency distribution of the number of shirts to order,* with the tally rewritten as the appropriate number of shirts. The increments in thinking are pursued further. Now it may be that this simple frequency distribution could carry additional information to make it more useful to you. Perhaps the total number of shirts would be useful, so that you can determine the cheque amount, that is, 100 shirts at $5 each is $500! Overhead 3 is an expanded table revealing the cumulative frequency distribution of the shirts.

We already use statistical ideas; although, we may need to attach a new name to what we call those ideas. Now, where are our cooks? Have you ever sliced up an apple pie? We have a family of five wanting to eat some of this pie. How will you cut up the pie? Collect two or three responses.

Students' response: The class will volunteer that the pie should be cut into different size portions.

Teacher's script: Dad likes a big helping, mum a medium piece, the three kids will fight if any one of them gets an extra crumb, so they'd better have pieces of equal size. How better to represent the apportionment of the pie to the five family members than with … *a pie chart*. Each piece we label in our heads as dad's, mum's and a piece for each of the kids. We might even go further to assign proportions or percentages to the shares allocated - because the kids will fight if anyone gets one more crumb than they!

Overhead 4: Portions of the pie for each family member



We could just as easily represent the carbohydrate, protein and sugar and other content of different food types using pie charts or describe some characteristic of the group of people who form our sample. (This is followed by a pie chart representing something of relevance to the group) And of course the pie chart would be better in rank (or Pareto) order. … well, from apple pie to weight watchers. Would someone like to describe what you do on that little piece of paper stuck up on the refrigerator… when you are on a diet?

Students' response: A picture on the refrigerator, but pushed, the question rephrased someone will suggest using a line chart (although not necessarily by that name).

Classroom scripts: You need no introduction to statistics!

Teacher's script: Some of us are familiar with drawing *line charts*. Reveal a line chart showing weight-loss over 20 weeks or so. Somebody will have used one - certainly they will recognise them … Well where are our photographers? You are at a school reunion, there are thirty students and you have one shot left. How will you take the photograph?

Students response: There may be several, but most will line the participants up in rows (*order*), tall to the back, short to the front and some sitting down....

Teacher's script: *Ordering the data* in a data set can also be a useful exercise in providing meaning to our observations. How can we order the data in Overhead 5 to have it convey more to us? What happens when we organise the data? What happens when we separate the males and females?

Overhead 5: Physiological data from a study on cardiovascular risk factors.
Unordered data

height cm	weight kilos	sex	serum cholesterol mm/Hg
170	65	M	6.2
171	80	M	7
172	74	M	5.8
164	59.5	F	7
…			
168	62.5	F	4.9

The next overhead shown is the data sorted by the sex of the patient and questions are asked: Can anyone see any patterns in the data now? Take a look at the weight column. Males seem to be heavier than the females. Males would appear taller although we have one very tall woman. A clear pattern with
cholesterol is not evident, although perhaps the males also tend to have higher cholesterol levels.
Teacher's review script: Now it is the teacher's turn to wrap-up. Always get all those ideas together at the end of a lecture.

In summary: statistics is, in part, about making sense out of data. The techniques we have reviewed tonight may be termed exploratory data analysis techniques or descriptive statistics. Some of the ones we have met?

Overhead 7: Statistical ideas

- ordering data
- tabulating data: tally sheets; frequency tables; cumulative frequency distributions
- graphical representations: pie chart; line chart
- measures of centre (mean, mode, median, trimmed mean - note how Olympic diving is scored)
- measures of variability (range,...)
- sampling. Our information is based on an appropriately (in some sense, representative) chosen sample
- the correctness of assumptions about the real world or the theoretical distribution are important.

As mums, dads, singles, mere mortals we are already using statistical processes, we are observing and making meaning of data. As scholars, undertaking university studies, hopefully you will come to appreciate the value of statistical techniques and expand your repertoire of statistical skills.

The second lecture was similar to the first in that students, after reviewing their knowledge of questions that are asked in statistics, were led to the question of relationships, trends and causality. Once again, a huge variety of statistical ideas was elicited, this time pertaining to the topic of analysis of bivariate data.

Activity-based exercises. The dominant mode of teaching involved students completing assigned exercises. These engaged the students in statistical thinking as they actively generated data. Students were introduced to a topic by *scaffolding* from old work to new work, by reviewing previous work and indicating where the current work was to fit into the overall framework. The exercises were designed so as to provide data that would lead students to appreciate how variability affects that which they observe and measure. These data are summarised and presented to (or by) the students in the class using graphs and tables. It was intended, if for no other reason than to keep the full-time workers and part-time students awake, that classes be kept lively and fun. Students participated in exercises on sampling, physical and psychological measurement and probability. Throughout all exercises, students' observations and their thoughts about the data were gathered and translated into statistical terms by the teacher.

One exercise involved calculating mean age for many different samples. Students were given verbal instructions to calculate the mean age for samples of size 2, 3, 4, 5, 6 and finally for the class population. They were also provided with the work-sheet (Exhibit 5.3), which provided a written description of a task. This description gave the transition from everyday English instructions as spoken, to written, mathematical notation. They were also given spoken instructions, omitting any reference to the mathematical notation.

The aims of the sampling exercise were:

1) to have students understand the notion of variation in sampling, that the sample provides an estimate of some feature of a population, and that as sample size increases that the estimates are less variable; and,

2) to familiarise students with the language of population, sample, estimation, variables, ways of sampling, measure and means.

Exhibit 5.3 Sampling exercise work-sheet

Work-sheet In today's exercise I want you to repeat the collection of some data, with each repetition involving a bigger sample of people.

1. Your age (X_1) _______
   other's age (X_2) _______
   Total of both ages _______
   Mean age $\overline{X}$ = Total of both ages ÷ Number of ages =

2. Your age (X_1) _______
   other's age (X_2) _______
   another's age (X_3) _______
   Sum of ages _______

   $\overline{X}$ = Sum of three ages /Number of ages =

3. Your age (X_1) _______
   other's age (X_2) _______
   another's age (X_3) _______
   another's age (X_4) _______
   $\sum_{i=1}^{4} X_i$ _______

   $\overline{X} = (\Sigma X_i)/N =$

4. (X_1) _______
   (X_2) _______
   (X_3) _______
   (X_4) _______
   (X_5) _______
   $\sum_{i=1}^{5} X_i$ _______

   $\overline{X} = (\Sigma X_i) / N =$

5. Repeat for a sample of six people.

6. Calculate the mean age for the entire class.

The process was completed by all students without difficulty (some practise on long division was required), except that some students were observed as anxious as they attempted to read and understand the sigma notation. They were told not to 'worry about the work-sheet' until the end of the procedure. The students were deliberately kept focussed on doing the task. The aim was to establish that they could do the work, creating confidence in their capability, before moving to examine the written, mathematical language. The statistical ideas and outcomes generated from this sampling exercise are presented in Exhibit 5.4.

Exhibit 5.4 Sampling exercise outcomes

Outcomes: Sampling Exercise

The means were recorded on the blackboard under the headings sample size 2, sample size 3 etc. The mean age of the entire class was calculated. The framework that was provided at the completion of the exercise was that this exercise was like a mini research project wherein data is collected only on part of the population of concern. Students were asked ‘what do you observe?’ The exercise generated discussion. The ideas and concepts, and the exercise itself were then placed in a framework namely, how we approach research.

Teacher's script included:
In gathering information about eg the population of Australian men and their level of cholesterol we take a sample in order to estimate something about the entire population. Here we took many samples to come up with an estimate for the population of Gateway students - this class. Generally the mean of the population is not known. In this instance, the Gateway student population could consist not only of this class, but the other two classes this session and perhaps all previous classes or even all future classes. A population can be a fixed and known group of people (objects,...) or a hypothetical group or perhaps even an abstract concept.

How to take a Sample
Classes gave ideas as to how they selected their samples eg one person included only persons older than himself, thus giving a poor estimate. Most selected those sitting close to themselves. Some chose those who were only young. Discussion of Morgan Gallop Poll, and other surveys eg being asked in the street, telephone surveys, quota systems, stratified sampling (choosing to represent all age groups) was identified. Other questions were asked. Would we stand outside Wollongong pub to survey people in order to find out the average consumption of alcohol in Australia? What would happen if we did? How would we sample if we wanted to investigate medical care in hospitals? Randomly select big hospitals, small hospitals, private hospitals, public hospitals, those with specialised function those without. Then who would we sample within a hospital?

Effect of increasing the sample size
By circling estimates furthest away from the means it was readily seen that as the samples got larger the estimates tended to get better. Not all were better. None was exactly the same as the population mean. Many were above and many below the population mean, taken for this purpose to be the mean for the class. Most of the estimates were close to the real mean, but a few were far away.

Mathematics
The exercise allowed practise in adding and dividing. Some misuse of mathematical language was noted. Where students expressed answers as integer part plus fractional they were required to express this as a decimal. Students recognised that they moved from a language they understood to one that was ‘alien’. Learning the language of Mathematics was likened to learning another language - it may take some time and repetition before the meanings of the words sink in, however the mathematical operations to be performed are understood. Using what they had done in the earlier section of work, they were, as a group, able to ascertain what the symbols meant.

'Getting it together' sessions. The start of each session generally involved asking students the purpose of the previous week's work and what we had found. This was partly for recall purposes but also to help establish a statistical framework through linking this week's work to that of the previous weeks. At the end of each teaching

session the framework was again evoked in summarising the lesson. After the major assignment, an entire session was spent reviewing key concepts and how they can be connected.

5.3.7 Data Collection: Statistics Understanding

The teacher's theory building is based on the incidents, student comments, discussion as they unfold in the classroom and supported by data generated through student assessment tasks.

As the teacher-researcher taught she recorded *critical incidents* in the classroom. The recordings of these incidents provide data regarding the phenomena which had an impact on teaching and the teacher's thoughts about the teaching process.

Students generated data regarding their understanding of statistics through completing a variety of assessable tasks throughout the session. These included:

- five statistics homeworks required students to reflect upon statistics exercises completed in class and to apply the concepts encountered in a piece of writing. Full marks were given for the submission of each of these homeworks. To gain feedback students needed to submit work. The feedback provided was directed at clarifying, and extending the statistical concepts identified by the student and to improve student writing. These homeworks led to a major piece of work.

- a major statistics assignment in which students were to identify and discuss the major statistical concepts encountered in class.

- a final examination question on *Pearson's correlation*. This was, considered to be both a mathematics and statistics question, with both calculation and interpretation of the *coefficient r* required.

- a final examination concept mapping question where the major concepts were to be identified and linked.

- a final examination question. The question was as follows: 'One aspect of statistics involves the communication of information. This we may do by numerical, graphical and tabular summaries. What general features of the data may we be interested in communicating and why?'

- a final examination question. The question was as follows: 'We collect data from samples in order to estimate something about a population (eg the mean). What features is a sample likely to have in order to provide a good estimate. Why is this so?'

5.3.8 Outcomes: Learning and Understanding Statistics

In this section the 'on the job analysis' or reflection-in-practice is presented. The data selected are those which had the greatest impact upon the teacher during the teaching session. These incidents guided the teaching process and ultimately the development of a grounded theory of teaching. The data are drawn from the critical incident analysis, major assignment and concept mapping examination question.

Analysis : Critical incidents

Some of the landmarks during the first six weeks for this component of the course are presented in Exhibit 5.5.

Exhibit 5.5 Critical incidents in teaching

Critical incidents in teaching	
1 Activity:	Mini lecture: You need no introduction to statistics.
Teacher:	Students were responsive and seemingly all engaged
2 *Activity:*	Sampling exercise. Find the mean age in samples of size 2-6 students. Ideas - Examining the process of taking samples to estimate the mean of a population .
Students:	Students were actively engaged in discussion and the generation of statistical ideas such as variability in estimates. The quality and quantity of concepts elicited was excellent. One week later, students were asked the purpose of the previous week's exercise. 'Long division' was the ready response! The appropriate response was also elicited.
Homework:	In class, you have met the following language and/or concepts. What do these terms mean? Show they relate to each other either through diagrams or words. To this a student responded: 'Why don't you give us a sheet with all the terms to learn? There is too much to know.'
Teacher:	Observed: (1) Some students were seen to concentrate on what proved difficult for them rather than the most important concepts. (2) Students responded with text book type definitions of terms rather than the meaning as they intuited from exercises. Students were asked to think through what it was they did in class and thus to identify and re-create the ideas, the meaning and then the appropriate statistical terms. Relating terms and ideas was poorly done.
3 *Activity:*	The physical measurement exercise required students to: weigh themselves 3 times; measure the perimeter of a desk 3 times; measure around the circumference of a given ankle. They were also to compare these measurements with estimates made prior to the first measurement of each.
Student:	Some students changed the task, making it more difficult eg using formula methods to determine the perimeter or by calculating means.
Teacher:	Discussion was fruitful once again yielding variability as a concept. All were actively engaged. Students were required to summarise and display to the group their findings. Discussion ensued
4 *Activity:*	Psychological measurement. What type of person are you?

Critical incidents in teaching	
Teacher:	Tasks continued to generate concepts (eg. reliability & validity).
5 *Activity:*	Statistical ideas.
Teacher:	This built on student's use of information to make decisions, draw comparisons and to display and manipulate data. They could identify as statistical thinkers.
6 Activity:	Providing a framework. My sense was that students were struggling. They needed help to integrate what they had encountered so far.

In week 7 the major assignment became the focus of attention. The students' success or otherwise in understanding the relationship between variability and other fundamental issues in statistics was evident in this assignment. In the assignments questions that are asked in statistics, regarding the nature of sampling, measurement, display of data and how decisions are made in statistics were all evident.

Analysis: major assignment

The major assignment due in week 7 asked:

> Throughout the course we have been identifying, using and reflecting upon a number of statistical concepts and ideas. Weave a story around as many of these concepts as you can to show the place of statistical concepts in building knowledge.

As with all assessment tasks, a range of responses was elicited. Whilst in a group discussion after the assignment, the group could identify that to tackle this problem they needed:

- to firstly identify what concepts they had met, then select those which were the major concepts;
- determine how they related to each other and how they related to building knowledge; and,
- only then would the writing commence, weaving examples around the various concepts.

Only a very few students could individually identify, aggregate and exemplify the concepts drawn from the various exercises we had undertaken.

What was observed early in the course was that students could be diverted from the major concepts by the concepts they found most difficult or by concepts they found easy. What was unexpected was that, at the completion of the final assignment this should still be the case. It appeared that the idea of selecting the most important or a good coverage of concepts is not automatic. Should this be surprising? Perhaps it should not. In the next course, this phenomenon was to be observed more closely.

Earlier work had suggested that students found it incredibly difficult to link the statistical concepts together, the automatic inclination of most students was to provide discrete definitions for each term. This too was evidenced in the final assignment. The impression formed was that students needed more experience in categorising, determining similarities and differences in ideas.

The major assignment also revealed interesting patterns in how students wrote about concepts and their relationships. For many students, the essay writing essentially involved naming concepts without exemplification and without establishing the relationship between concepts. An essay could read (in part) ‘In statistics, we have learned about many useful things, the mean, the mode and median.’ In a second type of script the statistical concepts were embedded in the script and exemplified in a way that conveyed meaning and showed the relationships between concepts. An example of writing that conveyed meaning is presented in Exhibit 5.6.

Exhibit 5.6 Sample of student writing

The collection and analysis of data is done by sampling the public or consumer, either by random selection or a calculated cross section. It is important that the sampling be done in a controlled manner, otherwise the sampling will distort (bias) the figures... In education and psychology statistics are used to measure and compare individuals performances in relation to the larger community.

Literacy and language teachers often read extracts from the work of students to other members of staff and from this I was aware of the quality of writing that was possible. From this cursory glimpse, it seemed that students did not necessarily transfer their writing skills to disciplines that they did not see as involving language. Some students who had presented sophisticated writing in one subject failed to do so in the statistics component. This was an issue that needed to be pursued.

The week after the assignments were submitted, the class worked through the process of eliciting the concepts covered during term, identifying the major concepts, associated concepts and the ways in which ideas could be linked. Their assignments were returned with both individual commentary and a set of group comments on the task (Exhibit 5.7). One week after reviewing the major assignment, students were tested again on their understanding of the main statistical concepts.

Exhibit 5.7 Teacher's comments on the major assignment

Comments on the major assignment

Throughout the course we have been identifying, using and reflecting upon a number of statistical concepts and ideas. Weave a story around as many of those concepts as you can to show the place of statistical concepts in building knowledge.
What are the key aspects of this question?

Preliminary work

- Identifying the statistical concepts and ideas encountered in class.
- Which are the major concepts and ideas & what other ideas cluster with them.
- How do these concepts relate to each other?
- How do these concepts relate to building knowledge?

<u>Writing the assignment</u>
The concepts and ideas are connected; they do not exist as isolated definitions, so weave them together to give a coherent presentation of the concepts. Show how the concepts relate to each other and how they relate to the overall building of knowledge. The strongest presentations tended to take the one theme rather than lots of different examples to illustrate the points.

****

STEP 1: Identify statistical concepts and ideas. Here are some that we met in the course:

• variability; inherent variability in data; variability in estimates due to sampling; variability in estimates due to measurement
• sampling; better samples; random samples; bigger samples; estimation about the population based on samples; populations; bias due to sampling;
• measurement; standardisation; calibration; measurement error; nominal; ordinal; interval measurement;
• variables;
• display of data; pie charts; line charts; tables to present data; tally charts; frequency distributions; cumulative frequency distributions; histograms; highlighting variability; reducing variability; plots of seasonal data; putting a smooth line through data; conveying meaning;
• statistics as estimates of the population parameter; measures of central tendency; means; medians; mode; measures of variability; range; minimum, maximum;
• statistics as tests; t test;
• questions; is there a relationship between one variable and another? causality; Pearson's r; is this score different to the mean of a group of scores?
is the mean of this group different to the mean of another group?
• decision making; forecasting; predicting; building knowledge
• probability models; assumptions; there are different probability models; comparing data to assumptions about the world in order to make decisions;
• brainstorming; theory; previous findings
Of course, listing terms in this manner does not display your understanding of them. That understanding is more evident when the ideas are drawn together and linked in some manner.

STEP 2: Which are the most major concepts and ideas - start to draw the concepts together, group like concepts with like.

• *Variability.* How do we answer questions about data given that there are many sources of variability in the data: variability due to sampling, measurement and the inherent nature of data.

Comments on the major assignment

- *Sampling variability.* The different samples that can be drawn from the population will provide different estimates of the population parameter (as we illustrated with our activity using beads, and finding the mean, range, median, mode of samples of beads).
- *Measurement variability.* Answering many questions about variables involves measurement. However, measurement also involves variability.
- *Questions asked.* Is there a difference between this group and another group? Is there a relationship between this variable and another variable? eg tobacco smoking and death from lung cancer or time and the blooming of flowers?
- *Displaying data.* Aims: to convey meaning through displaying the variability in data, showing changes and differences; to reduce the noise in the display of data, helping make sense of data by eliminating the effect of seasonal trends, lines of best fit put through points on a graph. Types of displays include: line charts, tally charts; ordering data etc
- *Decision making ; probability models of the data*

STEP 3: How do these concepts relate to each other. Try to map them, describing why or how the concepts are linked. Try to find examples to more clearly explain the concepts.

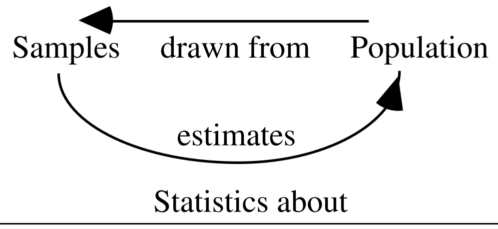
STEP 4: This involves weaving some sort of story around the concepts do as to illustrate there usage. The concepts their links and how they relate to building knowledge are the most important aspects of this assignment.

Analysis: examination question

The final exam, one week later, tackled this issue of the major concepts again, but required a mapping of the relationships between concepts. (Refer to Exhibit 5.8).

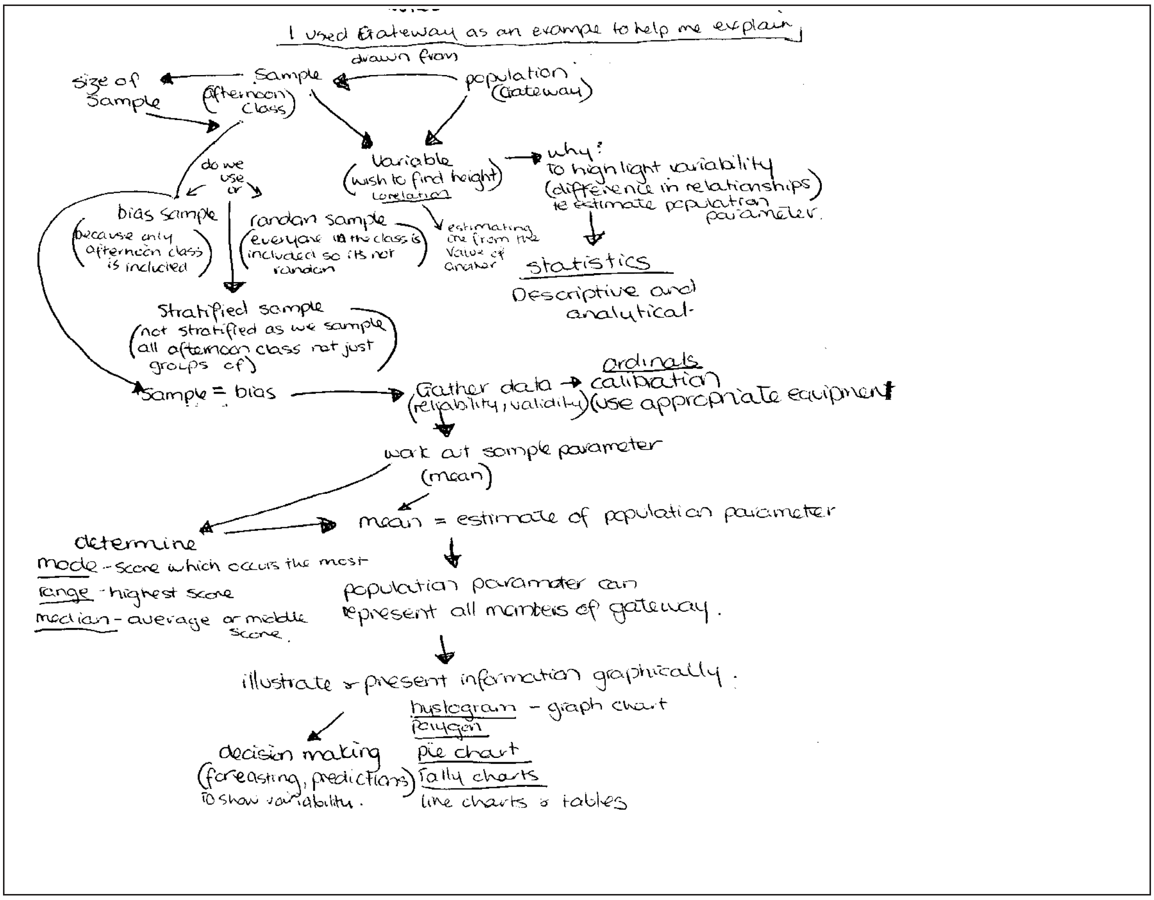
Exhibit 5.8 Concept mapping exam question

Draw a map, showing up to 15 of the most important statistical concepts that we have met in class and how those concepts may be linked together. One simple example of how the two concepts samples and populations may be linked is as follows:



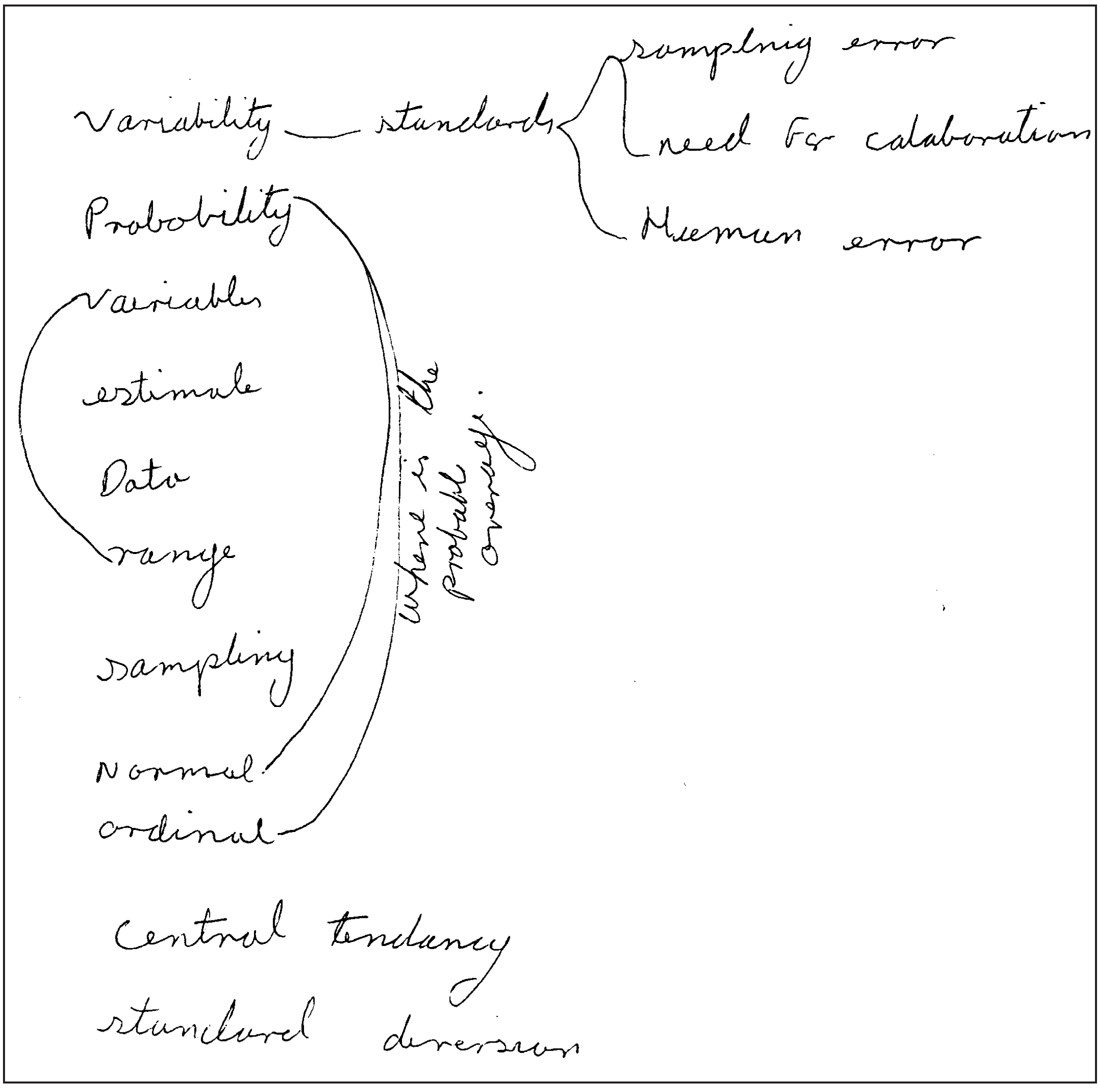
Of the 40 students who completed the question, only five students (12.5%) had variation as the central statistical concept, despite the fact that all units were directed at examining the relationship between variability and some other concept. An additional 14 students (35%) included variability as a concept. The maps could be sophisticated with linkages well established, as in Exhibit 5.9.

Exhibit 5.9 Student map of the major statistical concepts



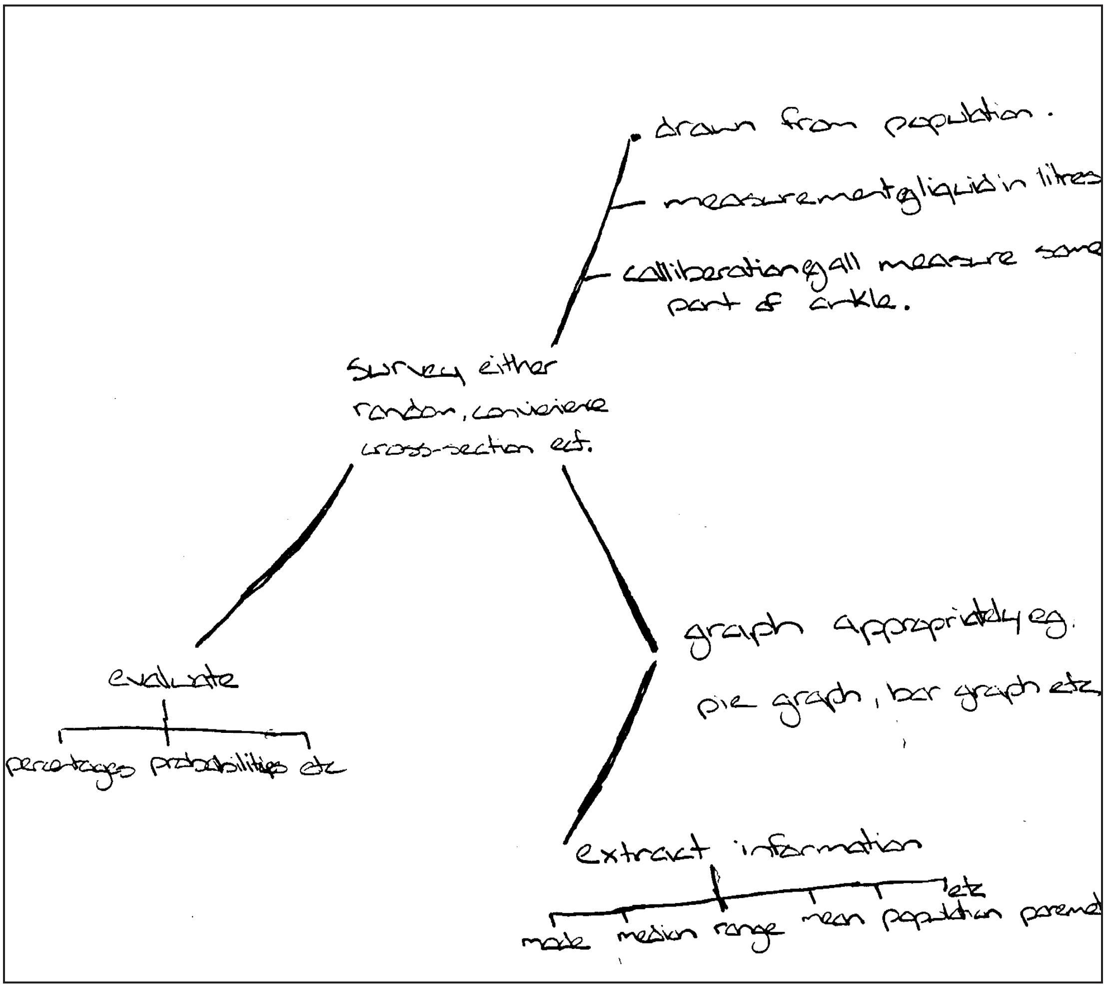
The maps could also be simplistic and lacking in linkages, as in Exhibit 5.10.

Exhibit 5.10 A concept map displaying difficulties in both the understanding and relating of statistical concepts.



Six students (13%) did not attempt the question and one student sat a supplementary examination that did not include this question. Twenty-five students (62.5%) did not include variability as a concept but focussed on the other major concepts. Once again, the maps could be sophisticated, with linkages and elaboration of associated concepts or relatively simplistic, as in Exhibit 5.11.

Exhibit 5.11 Concept map displaying clusters of statistical concepts.



Of the forty completed questions, variability was included as a concept by nineteen (47.5%) students contrasting with the inclusion for other concepts; questioning (n=12, 30%), sampling (n=35, 87.5%), measurement (n=30, 75%), description and analysis (n=33, 82.5%) and conclusions/decision making (n=21, 52.5%). Despite the fact that there was an explicit emphasis on variability rather than an implicit emphasis, as for example on description and analysis, there was less recognition of variability as a key concept. Design as a concept was included by only one student, not surprisingly perhaps, given there was no specific unit on the relationship between variability and design. Rare also was the presentation of a complete statistical process (ethics, questioning, design, measurement, sampling, description and analysis, and decision making). This is not surprising, as the comments returned to students after their major assignment also lacked the sense of an entire statistical/research process. Once again, this aspect was implicitly rather than explicitly drawn. To more firmly establish variability as the central statistical concept the curriculum needed to be modified. A unit on the important concept, *design*, was also needed to strengthen the awareness of a statistical process.

5.3.9 Overview of Implementation 1

Have you ever tried to describe you own teaching behaviour? What is it that you do? It is incredibly difficult. Take a step beyond describing what it is you do. Is there is a nice coherent theory to explain it all? As you begin to apply it, you find another nuance, another behaviour to explain. There are many aspects of the teaching, learning and assessment process that have not been examined closely, and many more that perhaps have not even identified. Those that have been the object of attention have been the aspects that were readily evident throughout the teaching and they are the aspects that have directed the changes in teaching. One major aspect of the teaching and learning process that required specific commentary was the role of language and experience in teaching statistics.

Overview: Language and experience in the statistics classroom

The statistical and mathematical literacy subject that was the focus of this implementation built upon an activity based approach to learning. Not only did students produce their own data but also they inspected, summarised and discussed the class data. They generated all of the statistical concepts and ideas utilised in the course. The teacher’s role was to focus and refine these concepts. Often, the role was to couch the students' ideas and language in formal statistical language. At other times it was to help link the ideas generated. Associated with these aims was the aim to establish *ourselves* as statistical and numerate thinkers, through identifying the statistical and mathematical ideas that are utilised in everyday life, and to extend these ideas so as to understand how they are utilised in academic practice. From the first week, homework required students to reflect upon the ideas that they had met in class and to identify how they used those ideas or similar ones in everyday life. Ideas generated were returned to frequently in later weeks in order to build up the picture of what statistics is about. This approach provides a contrast to that of many courses in which a topic is taught and then the next topic, without continually integrating the new ideas with the old. The ideas that were forthcoming in lectures and exercises were to be used and referred back to throughout the entire course. The exercises generated a large number of statistical ideas and concepts and, in general, high quality discussion.

Fundamental to the teaching was the usage of the language and ideas that the students, as a group, had generated in making observations about data. The language and ideas that students provided such as the 'youngest' and 'oldest' in class were translated to provide the statistical equivalent, the 'minimum' and 'maximum'. It was apparent after use of this teaching approach that:

- in the final assessment, most of the students were able to invoke many statistical concepts and demonstrate their meaning through mapping the relationships between concepts, although variability was not generally identified as the most central concept.
- that for some there were still difficulties with the language. One student's comment showed a lack of awareness that all the statistical language used was overlaid on ideas the students firstly presented in everyday language:

  > If the statistics could be worded in lay mans terms I think more people could grasp the concepts better. Then build up to the terminology. How that can be done, I don't know.

- many students were able to establish ‘themselves’ as statistical and numerate thinkers through identifying the statistical and mathematical ideas that they utilised in their everyday life. The following are three examples of comments made in the evaluation of the course:

  > The mathematics & statistics program’s strength was due to the identification of these subjects with day to day life. This broke down my barriers that I had towards Mathematics and Statistics.
  >
  > I feel that I have progressed forward as I had not a lot of knowledge in these areas. It has opened up my awareness of how often mathematics and statistics are used in everyday life and that they play an important part in our lives.
  >
  > The program has highlighted for me that the statistical knowledge I have and use everyday can be put to other applications.

- a small number of students was almost intractable in the belief that, mathematics and statistics are subjects they could never understand. An example of this occurred in the sampling exercise (Exhibit 5.2). Students, who had continued the process of calculating the mean and who understood what they were doing, found it difficult to accept that they could understand what the sigma notation meant.
- whilst students exceeded expectations in being able to articulate statistical ideas and concepts this did not easily transfer to writing about or mapping these concepts. Suggested course changes made by students include learning how to ‘write about mathematics and statistics’. For at least one student there was an even more basic discovery, namely that there is a relationship between language and mathematics and statistics. She comments:

> In the mathematics/statistics area I have learnt that Mathematics is connected with English. Up until the *Gateway* course I saw the two as completely different areas.

It was not possible for her to progress with either mathematics or statistics until this basic connection was made.

- under examination pressure some students reverted to an attitude characterised by 'I can't do it'. Several students who had completed their previous assignment which was also on this theme responded to the exam question with a blank page? They had taken part in a brainstorming session that elicited statistical concepts and determined which were the major concepts. The blank pages did not reflect a lack of academic knowledge. These students were able and generally articulate on statistical matters.

- as for the learning of mathematics, some students gained by finding strength within themselves to learn what was for them difficult:

> The strengths were that if you really wanted to achieve you had to do it on your own. I found, for me it was much better to work independently rather then to depend on others. My greatest weakness is confidence. I feel more confident now to know if I look up information that I will learn independently on my own.

To survive this subject, students needed to think back to the exercises and recreate the ideas they had previously generated through observation and understood in their own language. As this impression formed, so did another, that of the professional statistician, the professor-expert in my case study, who observes and seeks in data its many patterns, rather than mechanically applying 'the' statistical test.

As for the mathematics component, there are those who are now more enthusiastic about the statistics discipline than before (n=27, 65.8%), those for whom their enthusiasm remained the same (n=10, 24.4%) and those for whom enthusiasm declined (n=4, 9.8%) and others unknown or absent (6). When teaching students, it often seemed that they responded in accord with their needs and preferences not just to the subject matter and its delivery.

Overview: An invitation to learn

In the aftermath of teaching I was generally left pondering why some students blossomed and why others did not. In the search for answers, some readings seemed to connect (Purkey &Novak, 1984) and others did not. I found myself questioning:

> Did I, as a teacher, offer to each student an invitation to learn? Did I offer an invitation that could not be accepted by some? If I did not know the names of some of these students how well could they have received my invitation to learn?

> Whilst in class I am ever conscious of there being no wrong answers, just answers to the wrong questions. I used those answers as material to make discriminations and to further clarify ideas. I wondered if my feedback on assignments continued this theme or was it suggesting that the students were 'wrong'. Did the feedback promote the sense of an invitation to learn? The questions 'Did I invite my students to learn?' and 'What was the impact of my feedback?' were ones for which I had no answers. (Diary, 1994)

Overview: Metacognitive aspects of learning

Matlin (1994) interpreted metacognition as 'your knowledge of an awareness about cognitive processes' (p. 248). Metacognition is knowing, when you know and how best you come to know, it enables the selection of strategies to 'improve future cognitive performances.' At the end of the Statistical and Mathematical Literacy subject I was conscious of have taught a course wherein the content matter was *almost* incidental *(almos,t* as students were still required to display a certain level of competency at the end of the course in order to be deemed to have passed). At this stage my focus shifted to include an examination of ideas about metacognition.

Throughout the course, students could be considered to be learning to learn. They were to identify and use knowledge of themselves in order to learn their mathematics and statistics better. At the commencement of the course, students were asked to identify ways in which they liked to learn or could learn well. Throughout the course, students were asked to reflect upon: what they had learned; how statistical ideas were used in everyday life; and, how they were feeling and reacting during class tests.

From the student responses, it appeared that many students did not trust their ability to learn. They did not feel comfortable with the approach of reflecting, thinking and writing about what it was they had observed. Students would ask 'Why don't you just give us a list of definitions to learn?' This would have been a preferable approach for some. Students commented on the approach used:

> At the beginning, it didn't feel like we were going in any particular direction. (Although now I realised we were!) Maybe, if we were shown how these things fitted into the big picture of what we were trying to achieve, it would have lessened the panic of feeling a bit lost. Maybe, as each concept arose we could have fitted it into the framework from the beginning. I have learned how statistics and mathematics fits into everyday life, but that happened without me really seeing it until the end. I think it was a waste that I didn't know a bit more about the actual concept framework first. I feel my enthusiasm suffered for a while because of my confusion. It took away the fun.
>
> A little more teaching on the board would have helped.
>
> Try and map out first and then look deeper into the various statistical concepts we covered.

In this statistics component, rote learning was neither required, nor possible. Textbooks were not required. The aim was to have students learn from their observations. From the comments, students appeared to be better able to cope with learning through the structured mathematics program, even though they were primarily self taught and many had poor previous experiences with mathematics, than this approach to teaching and learning statistics.

Overview: Reflecting on collegial feedback

Details of the implementation of the first statistical and mathematical literacy subject reported in this chapter were reported in the papers, *A Statistical and Mathematical Literacy Pre-tertiary preparation course* (Porter, Griffiths, & Hedberg, 1994a), and *The role of experience and language in teaching statistical concepts* (Porter, Griffiths, & Hedberg, 1994b). These presentations allowed for feedback from colleagues. The sense that I gained from presenting and moving amongst mathematics and statistics educators was that the 'teaching' of mathematics, which focussed on how to perform a skill, drill and practice, was not one that would be considered acceptable, while the activity based approach to teaching statistics would be acceptable. Furthermore I rejected this approach to teaching mathematics, but the time constraints appeared to leave few options. In this instance the aim for mathematics was simply for students to be able to successfully perform skills such as calculating percentages. In statistics the students needed to demonstrate their understanding of topics often demonstrated through questions of the kind '*what might happen if...*?'. In statistics the calculation, say, of a Pearson's correlation coefficient r was only one of the lesser aims of teaching the topic. The substance in this topic was understanding the context of using Pearson's correlation coefficient r, how to use it and how to interpret it. I had anticipated that in a learning situation focussing on understanding rather than rote learning students would prefer a learning environment leading towards understanding. Incongruously, students felt more comfortable with the mathematics teaching than the statistics teaching, while I as teacher-researcher believed the approach used in teaching Statistics was better for providing students with understanding of what they were learning.

Overview: Reflecting in the context of a thesis.

The theme *Unpacking the expert in order to teach better* (Bain, 1990) was the catalyst which led to the re-packaging of the statistics curriculum and the attendant theme of *statistics is a study of variability*. The focus upon mapping concepts and showing how concepts related was based on the work *Learning how to Learn* by (Novak & Gowin, 1984). In this implementation, concept mapping highlighted the difficulty students experienced understanding statistical concepts. The approach of identifying and linking concepts was useful in highlighting the need for further changes to the curriculum and the sequencing and presentation of units.

There was room for improvement, earlier diagnostics, setting students on their self directed path earlier, simplifying and re-ordering some aspects of the curriculum, adding and replacing units in the curriculum, working on student writing and developing mapping skills as a way of honing students' understanding of statistical concepts. As the teacher, I needed to focus on developing students as people who can cope with the unknown in the academic setting.

5.4 Case Study 2: Revising the Literacy Curriculum

5.4.1 Introduction

In this section there is an examination of:

- the teacher reflections which led to a change in the curriculum for the second implementation;
- the nature of the changes in curriculum, with details of the additional statistical units taught;
- the theoretical basis for teaching and learning;
- classroom teaching and student participation through videotape;
- how students reflect and write about their classroom observations;
- a critical week, the low point in both teaching and learning;
- student challenges to be overcome in learning; and
- outcomes for the second implementation.

5.4.2 Teacher Reflections: preparing for the second implementation

> The voices of delight and the voices of protest are quiet. The rush, the multi stranding, the bits of paper, handouts, the marks, the missing assignments, who is coping? who is not? reviewing, framing the course, setting the tests is done. The exams are marked. The beads for the next practical have been found. The orchestration is complete. The course is over. The students have gone away. The examiners are soon to meet. It is time to rewrite the curriculum, time to frame it for the next class. (Diary, June 1994).

The course felt rushed. It felt as if there were too many bits and yet I wanted students to make sense out of the many bits, 'to put the jigsaw puzzle together' as one student said. A walk upon the beach provided a contrast to these feelings. Perhaps it is fanciful, but when I walk upon a beach, I feel that I connect with what is real; nature, stillness and quiet. I hear the many ponderings of my soul. Collecting seashells, with their many sizes, shapes and colours is all that is really important. It was the juxtaposition of these feelings that provoked further my sense of a need for authenticity in learning. Learning needs more than relevance. It requires authenticity. The title and description of a seminar, *The capacity to wonder,* triggered my thoughts (Diary, June 1994). The description of the seminar was as follows:

> All learning takes place in the domain of ignorance. Wonder, discovery, surprise and delight, invention, creativity, collaboration, fantasy, pondering, as well as error, failure, denial, and misinformation are part of the human experience of learning, of living: and their positive roles in education merit studied and playful exploration. (1994, Herdsa workshops).

The questions that I posed in the previous course were effective. Would you worry if little Tommy scores 2/10 in his Maths test? Why? In this instance students began to see that the decision-making they undertook in everyday life corresponded to a form of decision making in Statistics. The questions were extended to ‘Do students in this class perform differently to the students in the other *Gateway* classes?’ and ‘Is there a relationship between smoking and death from lung cancer?’ These are common statistical questions but trite in that the problems and the measurement were pre-defined. There was little wonder for students in these questions.

5.4.3 Curriculum Changes

After the initial implementation it was easy to see, not only the successes that had been obtained but, where additional units were necessary to help students build meaning, to better link concepts and ideas, and to facilitate students’ writing about them. Other units (eg the statistical process, questions of comparison, experimental design, and a reworked unit on measuring variability) were necessary additions to the units on sampling, measurement and probability. The inclusion of the units was facilitated, by three, additional one and half hour teaching sessions. These new units were to facilitate a greater cohesion between variability and the various stages in the statistical process.

The three additional exercises added to the Statistics Program originated as follows:

- *The statistical process (Seashell exercise).* Crystallising in my thoughts was an idea for an introductory practical. I thought to ask the students to generate all the questions they would want answered if they were to accept the results of an environmental impact statement. The environmental impact statement was to assess whether or not the sealife (as reflected by seashells) was different after a break-wall (or jetty) was built. (a fuller discussion of numerous issues arising from this is provided in section 5.4.6)

- *Comparison Questions (comparisons of sealife before and after).* For this exercise, students were given the two sets of shells, those collected before the breakwater was built and those collected after the breakwater. They were to draw a between the sealife before the jetty was built and after. The shells were used as an indicator of sealife. They were to present their findings in an appropriate form to the class.

- *Design exercise in order to examine the control of variability.* A conceptual gap surfaced in the curriculum during the first implementation. Whilst the students were made aware of various types of statistical questions, they were not made explicitly aware of the notions of design to control variability (or as subsequently differentiated, design to manipulate variability) and how this linked to answering statistical questions (refer, Appendix 5.8).

Three additional changes were made to the statistical units:

- *Measures of variability.* This exercise introduced students to measures of the centre of data and variability. In so doing students were to assess whether or not the measures were 'good' measures under a variety of circumstances. Previously students had learned these through the mathematics of sigma notation.

- The topic *Statistical ideas* was more clearly designated as analysis of bivariate data and it was described in terms of covariation. To strengthen the linkage to variability, the language used to describe correlation was that of, *two variables varying together*. Correlation chosen as a means of linking students to the formula oriented courses they were likely to undertake in 100 level university courses. It was to demonstrate that they could calculate what would be the most horrendous formula they would encounter. For students it was the ultimate achievement in the mathematics strand. Without this latter requirement, other more basic statistical units would have been included. However, as it was included, they were to demonstrate understanding of when and how to use the formula, and how to interpret it.

- The exercise *Psychological Measurement* (developing a questionnaire to predict those who would succeed in learning) was omitted as its focus was on learning issues rather than being central to the notion of understanding variability. Completing a *Dolly Girl* questionnaire questionnaire which generated ideas regarding reliability and validity was retained . All aspects of the physical measurement unit were retained. Recently Cox (1997) suggested that the issues of reliability and validity not only could, but should, be presented in terms of components of variance:

> The psychological metric notions of reproducibility, and of face, criterion and concept validity and of comparison with a gold standard are not normally presented in terms of components of variance but would probably be better done so. (p. 267).

I had not thought to do this, missing the opportunity to link all issues raised more tightly with variation.

The exercises in this implementation were more tightly focussed. In this implementation, the theme *Statistics is about variability* was extended to *Statistics is about variability throughout the research/statistical process*. Where the research process was loosely identified as a sequence involving ethical issues, incorporating appropriate expertise, questions asked, design, sampling, measurement, description, analysis, and decision making. Issues pertaining to variability may exist at each step in the process for example we design experiments to either control or manipulate unwanted sources of variation.

5.4.4 Clarification of theoretical basis for teaching and learning

Drawing on experience of the first implementation, the teacher felt more settled with the adopted teaching approach. Whilst some students were uncomfortable with the approach to learning statistics, the learning outcomes were pleasing, more-so than when the teacher had used traditional approaches to teaching statistics.

The pedagogy was more readily articulated as complying with a constructivist approach both in the development of and the teaching of the course. Constructivist approaches to learning contrast with transmission models of learning, wherein the student is basically seen to be an empty vessel into which the teacher will pour knowledge. As Barnes (1994) explained:

> People are not empty vessels to be filled with knowledge, but rational beings with active minds which construct meaning for themselves. We try to interpret data received from our senses, including information derived from social interactions, by conjecturing, predicting

> patterns, generalising and building theories...the statement also claims that knowledge is not contained in materials, whether text books or manipulatives, but must be (re)constructed by each individual for him- or herself. (p. 1).

Reflection is seen as an integral part of constructivism. Barnes continued saying 'engaging in activity of itself does not bring about learning' (p. 3). It is reflection that highlights the conflict between the old knowledge structure and the demands of new information to be accommodated within it:

> If we are reflective and inquiring, it is likely that we will encounter events which call into question our conceptualisations and we will be forced to reorganise our ideas. This reorganisation may require throwing out much of what we have constructed and reconstructing our schemes of knowledge. (Wheatley, 1991, p. 12).

These concepts of construction of one's own statistical knowledge and reflection were fundamental to all implementations of the Statistical and Mathematics Curriculum. During the second implementation the teacher was acutely aware that her pedagogical approach was rooted in these two ideas. The class exercises required students to carry out activities, to observe what they saw and to articulate and write about their observations. They were not given a package of information to take away and study. Through this method of teaching, they played a major role in the construction of their statistical knowledge. However, as I was to realise after the second implementation, students construct their version of what is presented in all teaching situations.

As with the first implementation, many students wanted the safety of being told what it was they were to know and what it was they were to give back to me. They also wanted lecture notes to define what it was that they had observed. Once again, the greatest difficulty for the teacher was not to give in to *the teacher's trap*, the desire to explain, but rather to play detective in firstly understanding the students' statistical ideas. Once the students' ideas were understood by the teacher, the second task was to find ways to further develop that understanding such that it incorporated the shared meanings of content and processes that are fundamental to the statistics discipline. In this context, tapping into the meaning that students have constructed was crucial. Within this framework, two aspects of students' work were constantly in focus:

- students' ability to establish and modify their statistical frameworks; and
- the role of communication between the teacher and students.

Teachers must constantly assess how it is that the students have constructed and re-constructed meanings. To establish what it is that students understand, or the meaning they have acquired, it is important that the students' learning environment is one wherein they cannot simply echo the teacher's or someone else's words.

Throughout the first implementation the teacher was an aware of the difficulties students had in learning and was attempting to build students' own awareness of their learning approaches. I reminded myself of my intentions through earlier writings:

> …This thesis is not about metacognition, in the final assessment the subject matter at hand needs to be mastered and is the focus of this thesis. It is about deep learning and the personal construction of knowledge. How can I help students such that their construction of statistics becomes more like that of experts? (Diary, June, 1994).

I was to again reflect upon this statement at the end of the second implementation, when the course name was deliberately changed from *Statistical and Mathematical Literacy* to *Learning Mathematics and Statistics*.

5.4.5 Teaching Methods: Video capture

As for the first implementation, the predominant teaching methodology involved structuring activities to generate discussion of what was observed. Students were to actively construct their own statistical knowledge. Important facets of the course were again the elicitation of ideas, refinement of them and acquisition of the appropriate terminology for those ideas. This approach was used, with partial success, to circumvent the attitude *I can't learn statistics*, for indeed the students were engaged in statistical thinking.

In the previous implementation the teacher was curious as to which students had a negative or positive reaction to the teaching. Would it be possible to identify these students from their engagement in class activities? Would it be possible to identify from the students' participation those who would not complete the course? In this implementation, permission was sought and gained from students to videotape the classroom sessions and to use these videotapes for research purposes. Whilst it was not possible to identify from the videotapes those who would experience difficulties, drop out, or not engage in the work, the tapes can be used to exemplify the pedagogical techniques used in the classroom and student responses to them.

The videos were collected during four of the statistical exercises (making comparisons, measurement, sampling and design). Three of these have been used to illustrate both how statistics is taught and how learning issues become evident in the classroom. The clips reflect the teaching techniques of all implementations of the statistical and mathematical literacy curricula discussed in this thesis. The student responses, their engagement with the tasks, the language used and ideas generated may also be considered to be similar across all implementations. The video clips have been included to overcome a problem identified by Olsen (1999) in examining video tapes of classroom teaching collected from Germany, Japan and the USA:

> …we don't have a shared vocabulary about teaching. What one teacher means by a term like problem solving is not necessarily what another teacher means and the problem is worse across cultures. (p. 30).

The videos have been edited to cut those students who did not want their images used and to produce short sequences demonstrating key features of the teacher and student activity in the classroom. They have been further edited to be short enough to be stored on CDROM. This second editing phase involves editing within sequences, removing for example the pauses as the teacher writes on the blackboard, between conversations or whilst a question is rephrased. What happens in the classroom takes a considerably longer than the clips portray. The clips and sequences have been chosen so as to illustrate both the teaching pedagogy and the generation and connecting of statistical ideas. Three units have been used to illustrate classroom activity. These are the units, *A Comparison of Sealife*, *Design*, and *Sampling*.

Statistics teaching: Images from the Comparison of Sealife Unit

The first statistical unit videotaped involves the comparison of the sealife before and after the building of the breakwater. Students are to use the seashells collected to decide if there is or is not a difference in sealife. In this unit, the feelings that may be evoked in students range from feelings of camaraderie to frustration and anger amongst students. Some students can dominate, some students might not be listened to, good suggestions can be ignored and poor ones put into action. The group dynamics and the feelings associated with them are important parts of learning and are evident in the video clip, *Comparison*. Evidence of students engaging in statistical thinking was also apparent throughout the unit. The three hour class is reduced to a seven minute clip as detailed in Exhibit 5.12 and viewable from the CDRom file, *Comparisons*.

Exhibit 5.12 Comparison clip

Components of the comparison clip

Introduction to the Comparison of Sealife Unit
The statistics exercise begins with the practicality of clearing the desks ready for the activity. The teacher then builds the scenario, drawing on ideas of design and collection of data from the previous week. Students are to find a way to 'measure' the two sets of data collected in order to make a comparison between the beach before and after the construction of a breakwater. Off screen, the teacher draws the students' attention to the overheads they had created last week to remind them of the questions they had come up with, the instructions provided in their workbook and other resources such as graph paper that they have at their disposal.

The groups begin work
Both groups immediately remove shells from the box. For the first group, all members are engaged in this initial action whilst in the second group the activity is more tenuous, with only a few removing shells. Discussion in both groups then intensifies as each group reassesses its course of action. The groups can hear each group's comments and decisions. The second group discusses taking a look at a handful from both groups, but decides to just look at their own group. An interesting moment is observed as a comment by one member of the second group leads the group in the activity on the exercise suggesting that the group look at a handful of shells. The group stands as the lead student stands. The decision as to what to do is still not made.

Mid-activity
After much discussion, the first group has decided to sort all the shells according to species and the activity and noise resumes. Meanwhile the second group accepts the reasoning of examining one quarter of the shells as being a fair representation of shells on the beach. Some of the off screen reasoning that the results would be as accurate as taking all is not correct. In the first group one member of the group recognises the need to name the shells so they can compare them with the other side. The group moves on to look at the classification names and the numbers of each species. The issue of whether shells 'are whole' and hence to be counted is raised. Activity and discussion continue.

A student leaves to investigates what the other group is doing
A student leaves the second group to investigate the activities of the first group. Another student clarifies that the purpose is to see if the breakwater caused environmental damage. Yet another student strikes a measurement issue and asks about the counting of a broken shell. Off screen the second group continues to spend their time debating what is required, eg what level of detail. There is a phase where this group appears quite despondent.

The student returns
The student returns and sits participating briefly.

A student tries to influence the group
When an opening occurs in the conversation the returned student suggests they compare with what the other group is doing. His comment does not change the activity, but he waits and tries again. The student suggests getting together with the other group. Initially, he makes little impact on the group (in contrast to the previous student, wherein the group also stood as the student stood).

Meanwhile the first group carries on
Meanwhile, the first group is progressing with the task (the returned student can be heard suggesting they groups get together). He waits (off screen) for another opportunity, but next time provides an explanation for the need to compare sets of shells. He engages the group in how they will make the comparison and what the conclusions would be if certain scenarios eventuate. Meanwhile the first group moves onto the issue of what to do with the broken shells. 'Are we going to give an approximation for the broken shells'.

Components of the comparison clip
The teacher watches and listens For much of this time the teacher simply observes. In this phase observation provides the material she will use later in this and later sessions. She draws on what students said and did, both in terms of statistical thinking, language used, decisions made and learning. The first group rate the species' condition of shells according to a scale of poor, fair and good. The student in the second group is still concerned that a comparison is not being made.
Teacher and student talk At this stage the teacher engages in discussion, supporting the student in his view about the need to compare the two sets of data. The student quite correctly suggests that the two groups have not got together to develop a uniform system for comparison.
A view of the group as the teacher talked to the student Still engaged in discussion with the student, we see the second view of his group. They have lost their animation. Following this discussion, several move to view the work of the first group.
Both groups now discuss the issues The second group moves to examine the work of the first. They engage in discussion about the issues associated with the comparison.

The lesson continues with students providing their data appropriately summarised and their decisions as to whether or not there is a difference in sealife.

Images from the Design Unit

In this clip the roles of the teacher and students can be readily identified. The teacher: introduces the tasks; draws together design issues and measurement issues as they arise amidst the student activity; draws attention to decisions students have made; introduces concepts; questions students; and integrates the work with previous units. Students work together in carrying out the activity, as they talk they reveal their ideas and language that the teacher can drawn upon in, they plan and explain what they are doing, they test the planes and measure, they question, listen, provide answers and make decisions. These facets of the teaching and learning process can be illustrated through the *Design* clip as detailed in Exhibit 5.13 and viewable from the CDRom file, *Design*.

Exhibit 5.13 Design Clip

Facets of teaching in the of the design unit
Teacher's role: The teacher provides an *introduction* to the experimental design task. The aim is to compare the distances flown by three planes.
Students' role: A few moments after the introduction students are *working* noisily.
Lecturer and students: *Interaction*: In the midst of the work a trio of students *interact* with the teacher. They are seeking *clarification* of the task.

Students' role: Elsewhere amidst the noise two students work quietly on planning the testing of the two planes.

Students' role: In the venue for testing the flight of the planes a *student explains* the design of their test procedure. The planes are tested. Measurement is required.

Teacher's role: drawing together the design issues. The camera focuses on the teacher and group, as the teacher *draws together* the design issues which are becoming apparent.

Teacher's role: following up measurement issues. The teacher *follows up* measurement issues that are becoming apparent.

Teacher's role: *Drawing attention* to decisions made. The teacher draws attention to the decisions which students have made. Language such as censored data has also been used and needs to be followed up.

Teacher's role: *Drawing attention* to the design decision to swap the order of testing of the planes.

Teachers role: *Questioning* whether or not there had been adequate testing.

Teacher's role: *Introducing* the concept of a pilot study.

Teacher's role: *Reviewing*. With the testing complete and a return to the classroom the teacher *reviews* ideas which have been drawn from the previous statistical units.

Teacher and students working with the data.

Teacher's role: reviewing concepts. The teacher *integrates* the design unit with other units. She focuses on the problem of making decisions about the difference of the seashells in the context of there being inherent variability in the number of shells that will be found on the beach at different times or on different days. This leads to the question 'How big does the difference have to be in order for us to conclude there really is a difference?'

Teacher and students' role: working with the data. Only an audio clip is available as students who had not given permission for use of their image were captured on both cameras. Data collected by the students is viewed on overhead. The student asks "what if we had only collected one data point for each plane" and similar questions to illustrate what happens when we sample data. Again, the issue as to making comparisons in the context of variability is raised by the teacher. A student suggests, again revealing statistical thinking, 'Wouldn't it be better to limit the number of variables?' leading to the role of a pilot study.

Images from the Sampling Unit

The sampling unit reveals the approach to teaching both statistics and mathematics and how the two may be integrated. As for the first implementation the 'teaching' of mathematics involved providing students with workbooks, which they worked through. Tests were completed in class to help them assess their rate of progress. After the tests the questions were completed on the whiteboard. The other source of mathematics instruction was incidental to the Statistics teaching and is also demonstrated in this unit as students encounter sigma notation for the first time. Throughout these clips, and off screen, the issue of students' statistical understanding and writing is evident as they progress through the activity. What is curious, given students preference for the mathematics rather than the statistics, is that the lack of understanding is principally in regard to the mathematical material. The clip, *Sampling*, viewable on CDRom portrays the transitions in activities and focus within the classroom to another as described in Exhibit 5.14.

Exhibit 5.14 Sampling Clip

Components of the sampling clip

Sampling Unit: beginning quietly as students finish a mathematics test
The opening segment captures a transition from a mathematics test to completing solutions. The teacher, off screen, is returning work to students. The quiet is a contrast to the noise associated with completing activities. Further shots reveal the classroom and students as the test completes and moves into checking work. A student clarifies a solution. Between clips another requests 'do that one again.' Another student compares work with that of a peer. As the students finish marking the teacher asks students to complete their diagnostic sheet (refer to the earlier, Table 5.1).

Reviewing previous statistical ideas
This segment was taken at the end of a 15-20 minute review period where the various phases of the statistical/research process has been elicited from students. This phase involves *scaffolding* from the previous work to the new topic of sampling and it concludes with the suggestion that sampling, which the students are about investigate is another source of variability.

Discussion regarding understanding and writing
Immediately after the introduction as the teacher moves to the centre of the classroom a student exclaims 'So what you are saying is that you think we understand what we are doing but we just can't write about it. Did anybody get that!…' Another student enters the discussion on writing. The final reference is to a discussion, prompted by the students through a request to the Head of Gateway. They were concerned that they did not understand the mathematics. The classroom discussion revealed that understanding the Statistics was actually their concern. However, when students were asked to describe the activities and what they had observed they recreated all that been covered. They understood seemingly too easily, it was as if there had to be this profound and as yet undiscovered thing that they did not understand. This was one of the many dialogues that took place on writing, learning and understanding.

Introduction to the sampling exercise
The sampling task is introduced. Each student is given a uniquely labelled card (with X_1, $X_2 \ldots X_{26}$). (They also used the card later when sigma notation is introduced as a mathematical exercise.) They are given the verbal instructions to calculate the mean age of samples, of size 2, size 3 …up to size 6, taken from the class. Students' workbooks specify the task as illustrated earlier in Exhibit 5.2. Within moments, whilst the teacher prepares the electronic whiteboard for the students' data, the noise level escalates as students ask the others for their ages and calculate the average age of the samples. For some, there was a need to address the issue of division with decimal remainders and this is dealt with either by the teacher or nearby student. On consulting the formal statement of the task students begin to ask questions about the X's and sigma notation, but at this stage they are told to 'just do it'. The time was not yet ready for the class lesson on sigma notation. As students calculate the mean ages for the samples and enter them on the board, the teacher is involved in a dialogue with a student who has requested help. The teacher suggests 'If you understood…', then the student responds 'when it comes to doing it with Maths or having to write about it I can't understand what you want'. In the background another student has discovered the 'hieroglyphics'. This is a typical scenario, the student can do the task (and the calculations involved in this instance) but cannot understand the mathematics of the task. It is for this reason that the teacher holds off providing the explanation of sigma notation. The mathematical language of the task is discussed when the students demonstrate that they can do the task and have gained confidence through completing the task. While many students enter their results on the board,

Components of the sampling clip
others work on. In this clip, a student checks with a classmate about her long division. Other conversations omitted were about 'how they were going' in general or specifically in this or other Gateway subjects. With the task completed, some of the students turned their attention to trying to decipher the meaning of the sigma notation.   *Generating ideas from data*   In this section the data gathering is completed and the means of the samples entered in columns from sample size 2 through to sample size 6. The teacher asks the students what they observe about the data. This clip records the investigation of a suggestion made by a student, that the means of the larger samples had a smaller range than the means of the smaller samples. By circling all means more than 4 years from the population mean students were able to verify this result for themselves. Thus they draw the first of their conclusions about desirable sample sizes. The issue of bias in samples is provoked by suggestions for taking a sample to study drinking habits and a question is posed as to how the Morgan gallop polls might be conducted. At this point a student takes the floor as he shares experiences of reading the results of polls. Another student contributes her experiences of having been a pollster. Others share their concerns about the possibility of collecting biased samples. The teacher converses with students, asking how students would take a sample of older persons and then responds to their comments and criticisms. The conversation leads back to tightening the concepts of cross-sectional and random sample.   *Tightening ideas and moving to mathematics.*   In this section of the tape the teacher is seen wrapping up the sampling unit by reference back to nherent variability and controlling variation through design. Then we have the separation of the statistical ideas from the mathematical. The teacher looks at the issue of being able to understand and do the task. A student is observed indicating a lack of understanding. Another student correctly interprets the student as not understanding the mathematics but as understanding the statistical ideas. This difference between understanding and doing the process and not understanding the mathematical language is again emphasised as the move to doing the mathematics takes place. The mathematics lesson follows by firstly drawing the analogy between learning sigma notation and learning the French language. The teacher conveys an understanding of the difficulty in learning the mathematics language, commenting upon her own poor language learning skills. We then shift to examine the meaning of the symbols for the mean and X_i as substitution of values into X_i is explained. The lesson continued with the introduction of sigma notation and reading sigma notation. We view again from when a student asks why we have a mathematical language and the response is in terms of it being a mechanism for abbreviation. Moving on to an example the teacher asks 'So what am I saying' and the students are to respond with the meaning of the mathematical expression and the teacher write up the terms as provided. Finishing there a few 'puzzled looks' and there is further clarification. There are more examples with a complete and tight reading and substitution into the formula. After this clip the lesson continues. Still there is a questioning as to whom is 'puzzled'. Still one student is puzzled and now this student's own X_i value is called into play. Working with her personal data and working with sigma notation continues until the end of class.

From a mathematics perspective, students continued to complete sigma notation examples each week. In this subject, the ultimate computation using sigma notation is the calculation of a *Pearson's r correlation coefficient*. The unit on relationships stepped outside the theme variability to some extent. It was used as a link between the mathematics of this subject and the most difficult of the mathematics the students would encounter in 100 level subjects at university, and this they are told.

Missing Images

One change in pedagogy that cannot be observed in the videotapes is the implementation of a period of reflection in the last fifteen to twenty minutes of several early units. Time was set aside for students to 'reflect' upon what they had done in order to provide a summary of ideas or to attempt to reorganise material such that similar ideas were grouped together. This change was implemented because of the growing awareness that constructivism placed considerable emphasis upon reflection (and in an attempt to reduce the out of class demands placed upon students). Homework was again directed at reflecting upon statistical ideas as they were utilised in daily life and, where possible, extending the list of statistical ideas they used.

5.4.6 Analysis: Students' reflective homeworks

In this implementation, the teacher was focussed from the beginning upon understanding the nature and source of students' reflections. Focussing on the process of reflection has generated many questions. Reflecting, it would seem, is a process that is thwart with difficulties, as the homeworks readily demonstrated. In this section, an analysis of the first two reflective homeworks is undertaken.

Homework: Mini lecture - You need no introduction to statistics!

The first class involved the presentation of a mini lecture, *You need no introduction to statistics!* (refer to the earlier exhibit 5.1). The homework was as follows:

> In this lecture you were reminded of your usage of a number of statistical ideas or concepts. What were they? Extend this list with other statistical concepts or ideas.

In an attempt to provide exemplars, and at the same time to give useful feedback to students, sample homeworks and associated analyses of them were placed on the student noticeboard. A similar type of analysis to that reported in this section was provided to students.

To maintain the integrity of the students, in addition to providing regular feedback to individual students, the approach of providing meta-levels of explanation was adopted. I attempted to answer the questions: 'How were these particular types of answers generated?' and 'What patterns were emerging in the answers given by all students?' In order to improve writing and to alter the direction of teaching it was considered that students needed to know what they were communicating to me about their understanding of the topic.

The crucial role of communication from teacher to student quickly became apparent. Perhaps a fault of the teacher's communication, possibly the term list should not be used in the same context as asking students to elaborate their ideas because for some, the writings were lists without conveying meaning. Students' reflections, at least as written, yield diverse interpretations of the statistical units. Even so, the patterns forthcoming lead me to think that there are characteristic ways in which students will respond, depending upon what level of meaning had been responded to (as subsequent analysis will suggest).

For the student Sal[3], the response to the first homework was a list. Whilst it was possible for the teacher to see that the student attended to much of the lecture it was not clear what meaning had been ascribed to each of the terms. His homework was as follows:

- tally score - tshirts
- dividing equal parts
- cutting pie - 1/2 - 1/3rds
- ratio - in proportion
- mode -most common
- median - middle range
- average - mean
- data- information collected
- trimmed mean - toss out the highest and lowest score and average the rest of the scores
- decisions (Sal, student spring 1994)

For Lyn the list was shorter and the last two items were of a questionable nature. The paucity and quality of content suggested a lack of active listening. The term common denominator suggested she had failed to distinguish between the statistics lecture and earlier mathematics work undertaken in class:

- frequency
- variables
- trimming mean
- pie chart
- validity of variables
- common denominator (Lyn, student spring 1994)

Another response to the group exercise was encouraging from another vantage point. A student not only listed terms, but did so in a manner that suggested a greater awareness of what they had heard. In using the term distribution, this student had linked two terms, pie chart and cumulative frequency, which were presented in different contexts. This student perceived at least two levels of meaning in the terms when he wrote:

> Centre
> mean, mode, average
> Distribution
> pie chart, cumulative frequency
> Comparison
> table, chart (David, student spring 1994)

These levels of meaning can easily be extended to a third (definitions of the statistical terms, mean and mode) and fourth which is *how to do it*! This short, incomplete, list is in some ways preferable to the more complete listing that follows. This listing closely parallels the summary overhead. It echoes too closely the lecture summary overhead (refer, exhibit 5.1, overhead 7) for it to be considered other than an echo. The student wrote:

- measures of central tendency
  mean (average), mode (most), median, trimmed mean
- variability (range)
- tabulating data
  tally sheets (shirt order), frequency tables
- graphical representation - display information
  pie charts, line charts
- ordering numbers in data
  (Additional to summary overhead suggesting active listening)
  - Decisions
  - Assumptions (Mary, student spring 1994)

Max is one of many students who deliberately sought out a text and echoed a list of terms rather than actually reflecting upon the class activities. He wrote:

> I borrowed basic statistics 3rd edition (Spatz and Johnston) and learnt about descriptive and inferential stats, parameters, population, sample, subsample, variables, quantitative variables, qualitative variables, nominal scales, ordinal, interval and ratio scales, as well as dependent, independent, extraneous variables and lower and upper limits, (qualitative variable which use the nominal scale). These are now part of my memory yet, I still do not understand their place. JIGSAW. (Dorris, student spring 1994)

Seeking out a textbook is not an uncommon response. Several other students found a text and used it for their homework. Whilst this is a desirable attribute in many courses, the task in this course has been to think through what has been undertaken in class and to extend to daily life. At this stage, those who sought a text conveyed little of what had transpired in class, and did not extend the ideas to daily life. The texts were not used to

---

[3]All names are pseudonyms to protect the identities of the students concerned.

further expand ideas covered, but moved to ideas not experienced in the exercises and not necessarily experienced in daily life. The act of finding a text was somewhat more time consuming than simply writing about the exercise in which he had just been a participant. From a teacher's perspective, I had an industrious student, but I could say little about what meaning was constructed from the lecture that had been delivered.

Homework: statistical process - questions to ask

The second week's reflection followed the unit on the *Statistical process* - *questions we need to ask.* The activity involved students in generating questions that were later classified as belonging to one or other aspect of the research process (or statistical process). The exercise is described in Exhibit 5.15.

Exhibit 5.15 Exercise to generate the statistical process

Exercise to generate the statistical process

A lobby group has called for an environmental impact statement regarding the effects of a temporary breakwater constructed in a small coastal inlet south of Ulladulla. They are concerned as to the possible impact this breakwater might have on sealife (as reflected by seashells left upon the seashore). The question that the impact statement is to answer is 'Is there essentially any difference in the seashells now left upon the seashore to those left upon the shore before the breakwater was built ?' Samples of the shells before and after the breakwater have been provided for your perusal.

Task.
1. In a small group of four or five, determine what questions you need answered before you can answer the questions that is required for the impact statement?
2. How can the questions you have asked be categorised or grouped in some manner?

The generation of these questions created a positive aura about this approach to teaching. The questions were illustrative of good statistical thinking yet, they were generated by pre-tertiary students. The questions reflected the sorts of issues that many senior students and researchers often fail to address.

The questions raised by virtually all groups of students tapped the issues that researchers needed to address in order to answer this question. The questions generated were later grouped by the class into similar categories. These categories were ordered and described as the research or statistical process. Questions were grouped in a similar manner to that which follows:

1. ethical issues. Who has commissioned this study and why? Ideas about deliberate bias were raised.

2. the nature of the real issue. Will the seashells reflect sealife? Or is the study flawed by this mismatch?

3. design and control issues. In the planning phase, the questions asked were numerous. Should we collect in the same season? How long after the breakwater was built should we collect data? Should we collect shells in different weather conditions?

4. sampling issues. When the design issues are settled data collection must proceed. From what areas of the beach and on how many occasions should the shells be collected. How should they be selected?

5. measurement issues. How many shells are on the beach before and after the building of the breakwater?…How many species of shells can be found before and after?…Have the shells deteriorated since the breakwater was built?…Has the colour of the shells altered?…Has the size of the shells changed? These questions also give rise to the notion of the different measurement scales. The nominal scale for naming varieties, the ordinal scales for ordering the state of deterioration of shells and continuous scales (and specifically ratio scales, where an infinitely divisible scale to measure the length of a shell or the weight of a shell can be used).

6. decision issues. Shells on beaches are inherently variable by nature. Given the seashells on the beach will never be the same two days running, how different will the difference in shells have to be before we say there is a real difference in shells.

The set of formal, written instructions for the homework for the second week is in Exhibit 5.16.

Exhibit 5.16 Homework question - generating the statistical process

Over the next week take note of the statistical ideas and concepts you are using that you met in the Unit on Seashells. Write a short summary of these instances and how the statistical ideas were used (or failed to be used).

Verbal instructions were also given. These were more informal and simply asked the student to think back over the exercise, identify the major ideas and to write them up, exemplifying the ideas where possible with examples from everyday life.

Student feedback included both a commentary on individual work together with exemplification of how the question may have been construed and written about. The process of maintaining students' integrity by making meta-level comments on student homeworks again revealed issues not just about statistical understanding but issues related to student learning.

Perhaps it was a false belief on my part, but I could not accept that these students did not know something of what transpired in a two hour activity based session. As Mai says regarding her homework:

> No notes for reference and no memory of the particulars to draw upon.
> (Mai, student spring 1994)

I accept that students cannot remember the nuances. However, when one recreates with students what happened during the previous week's exercise, knowledge of what transpired is evident. When the process involves reconstructing rather than trying to remember the activities, ideas are forthcoming.

The concept of levels of meaning is useful in examining the student homework in Exhibit 5.17. The student has completed the homework as the task was originally undertaken, that is, she has identified the questions she would want answered. This chronologically was the first step in the class exercise. In class, there was a further step where the questions were grouped according to a statistical or research process. For this student, and several others, the homeworks dealt only with questioning.

Exhibit 5.17 Student homework - question but no process

Student homework and teacher commentary

Students' homework

I have realised now that I use statistics in my work everyday. When an invoice comes in I have to ask many questions[4] before I can pay it. The questions I need to ask are:
Has an order been done for the goods?
Does the invoice match the order?
Have the goods been received?
All of the answers have to be collated and then the invoice can be paid or not paid ***(ie a decision is made)*** [5]depending upon the answers.

Teacher's comments

This homework can be used to illustrate the notion of levels of meaning that were illustrated in the previous week's work. Quite correctly the writer has tapped into the idea that there are questions that must be answered in order to make decisions. The student has applied this to an everyday situation as requested. However, the seashell exercise went beyond asking questions to classifying those questions into distinct categories. It is the application of each of these types of questions (eg measurement, design, sampling, ethics etc) that is the real guts of the question.

---

[4]The teacher has made bold the terminology of interest from the lesson.

[5]The student has made a decision but not explicitly labelled it as an act which is of importance in the statistical process, so the bold comment was added to the students work, drawing explicit attention to the statistical nature of the idea.

The reconstruction of what took place in class was the first and only phase in writing for many. Several students focussed only on the level of generating questions and perhaps applying those questions to an area of their everyday life. They omitted the explicit identification of the statistical/research process.

Another possible account of why homework was not completed to expectation could be that students might not be able to unpack questions in order to identify the issues which needed to be addressed. To undertake the homework, students needed firstly to identify the major ideas or issues that were covered in the unit. This reconstruction could initially be chronological but it ultimately needed to be a logical reorganisation of ideas. The logical ordering could most easily be developed in accord with the process identified. Then students generalise the ideas to issues found in everyday life. Ideally, the example chosen incorporated as many of the issues as was possible, rather than having a different example for each different idea encountered in class. This next example, Exhibit 5.18 tackled both aspects of the homework quite well.

Exhibit 5.18 Student homework - process and exemplification

Process and exemplification

Student' homework

Summary Unit on Seashells
During the session on seashells we were given the basic details regarding an environmental impact study being done on the effects of the construction of a breakwater and the effects this breakwater would have on the shore and seashells.

Main comments raised were:
* To collect samples of particular shells from various sections of the beach and collect data eg size, condition, age, type of shell and where on the beach the samples were collected.
* To ensure the ‘before’ and ‘after’ studies of the shore and shells were going to be carried out under the same conditions ie same weather, tidal conditions, swell etc.
* Why was the environmental impact being done in the first place and by whom?
* Why was the breakwater being constructed?
* How would the study be conducted eg. how would it be decided exactly what areas in particular would be studied eg. many small sections of the beach or larger sections of the beach - which side of the breakwater would be most effected after construction?

The various groups in class discussed what they felt was relevant in collecting all the data for the impact study. Some were suspicious of why the impact study was being held and who was doing it. The purpose of the breakwater in the first place was also raised, and in group discussion it became clear that while some points were important with some groups, they weren’t with others. We finished by discussing how we could represent the data statistically. We were able to study particular shells to note what we thought may be particular to that shell eg age, condition, size and location. Finally we discussed how we could use all the information we thought was relevant and present this in a statistical way.

I am using approaches similar to the way we discussed using statistics in my work environment - gathering information to use for an overall picture - we are currently receiving survey forms from

Process and exemplification

students which indicate what subjects they enjoy most, how they found the University environment, how they rated the teaching of their subjects, did they have any negative feedback and why they chose to study at this University. This information is being used to give feedback in relation to areas which may need improving and those areas which are attracting students to campus and assist in the marketing process.

Teacher's comments

If this homework is compared to the summary that I as teacher provided as to the nature of ideas covered in class, you will see that this student has covered most of them. Note that the purpose of the writing has been provided ie we have an introduction. Her writing includes a listing of the major ideas covered. That listing of ideas is placed in an overall context which provides it with meaning. Many of the items on the list are exemplified, allowing the reader to understand the meaning of the items. Note also that the second aspect of the question, the usage of these ideas in daily life has been addressed. This second aspect could be more fully elaborated by showing how each of the issues has been addressed. For example the issue of how students are sampled could be discussed. Design issues such as when the students are sampled could also be included. If I were to make a finer level of analysis of this piece, I would look more closely at each point. For example, I would suggest that the first point mixes two ideas, how the shells are measured and, how a sample of the shells is taken. However, these finer distinctions will emerge as the course proceeds. Overall, this piece of work tackles both aspects of the question very well.

The homeworks also suggested that students could make 'powerful use' of their statistical ideas (refer Exhibit 5.19). This student was able to implicitly demonstrate the use of the statistical process, without reference to the original exercise, but through reference to an example of his own creation.

Exhibit 5.19 Student homework - transfer of ideas

Student homework

I held the view that there were more job vacancies advertised on Wednesday and Saturdays in the Illawarra Mercury.

1. I counted the number of vacancies in Saturday's and Wednesday's Mercury. Though initially the figures for these two days was higher (Saturday being the highest) than those placed in Tuesdays, Thursdays and Fridays edition this was before other variables were taken into consideration.
2. When counting the advertisements, the ones that were run for more than one day eg an advertisement ran for two or three consecutive days was not counted on the second or third day of publication.
3. Categorising. The positions advertised were placed into categories determined by the hours or length employed.
   1. Permanent (35 hours per week or more)
   2. Casual (less than 24 hours per week)
   3. Temporary (less than two months)
   4. Seasonal (less than three months)
4. What type of employment was advertised eg professional, semi-professional, skilled and unskilled. Was there any relationship to the days they appeared?
5. What was newspaper circulation on particular days such as Wednesdays and Saturdays opposed to other days of the week. If so were there more readers of advertisements?
6. What percentage of employment adds appeared that only paid commission (are they real

Student homework	
	jobs?)
7.	What adds appeared consistently week in week out? eg letter box deliveries.
8.	What percentage of jobs was placed by employment agencies as a draw card (the jobs may not exist?)
9.	What percentage of jobs advertised paid no commission or no wage eg acting. (that is were not really jobs)
10.	What percentage of jobs that appeared in the paper had already been filled before the paper had been published.
Using a standard count of positions vacant does not necessarily give an accurate answer to the question but by including factors previously stated and allowing for deviations the initial statement may not necessarily be true. Even if this survey had been undertaken it would still be inaccurate as it would have only been taken over a 7 day, seven publication issue, which would discount factors such as seasonality. Thus it would not be a conclusive or true measure unless it was undertaken for a longer period of time.	

This student had been able to identify and use the many ideas that had been elicited in the exercise. In his analysis of the situation, he had fully pursued the implications of operationalising his definitions in various forms, and understood the implications of his sampling and design plans. He was also able to convey this without needing to refer to the previous class example - he had successfully transferred the ideas to a different topic.

Continuing with this form an examination and commentary on the many homeworks the teacher formed the impression that:

- some students were attempting to recall or remember ideas rather than to recreate or reconstruct the ideas from the prompt or exercise they had been given;
- many students were unable to unpack questions, to undertake both parts of a question;
- some students were unable to identify the major issues, they were often diverted by easy or hard issues;
- most students could readily talk about ideas but had great difficulty in writing about them;
- regrouping ideas in other than chronological order to provide a coherent and meaningful structure is a non trivial task that may require several attempts;
- students are not always aware of different levels of meaning.

Whilst it is easy to conceive of children being locked into some developmental stage where the ability to reflect is not possible, it was not really a difficulty that was expected for adult learners. In attempting to detect how students *were diverted from* *writing what was considered an appropriate answer*, it became apparent that reflection is not necessarily easy for adults. This observation stimulated further searching of the literature for discussions as to the nature of reflection in adults.

5.4.7 A critical week: Comparing the Seashells on the Seashore

After the third exercise (and third week) the teacher-researcher was confronted with student complaints about her teaching. She recorded her concerns in a diary. The extract is presented in Exhibit 5.20.

Exhibit 5.20 Diary recording of a critical week

This is the lowest point in the teaching calendar. The same time as last term. A number of students in one class have complained to the Head about the teacher in relation to mathematics. The blame it seems is at my feet. I am dismissive of them. This is communication that I can scarce but heed. Whilst I can discuss the issue with the students, I find at the moment a certain obstinacy in my approach and a certain disappointment.

Obstinacy: The mathematics component has been selected as the section wherein the students are to learn to be responsible for their own learning. This is within the capabilities of all the students. Certainly it is within the capabilities of prospective university students. This does not mean that students are left on their own, but rather that the onus of pacing and structuring their mathematics learning is left with them. The work is Year 7 work, the mathematics starts at the level of ‘write the number seven hundred and thirty eight in numerals’. For students to communicate that they need help they must firstly commence work, working through the booklets and in so doing identify the sections for which they need help. By submitting the booklets as they do them, I am able to comment upon them, write an explanation or find other explanations, that is I can communicate to each individual about their maths. At this point it is appropriate for students to ask for and receive help, as indeed a number have done.

What I am listening for, and what I am already hearing is an increasing number of voices. They are saying ‘I feel more confident, I could do my son’s algebra homework'. ‘How many places do you want the long division to be done to, I’ve got it out to seven places’. It was only last night before class that a student sought ten minutes help. Another said ‘I keep telling myself it is week 3, I am doing OK with this, I was never good at maths’. There is a certain level of pride expressed each time a student says ‘I think I’ve got it, the division is OK, its just the long multiplication now’. Pride as each small step is chalked off and the next step identified.

It is easy to step in and give a mathematics lesson. It is too easy for the credit for learning or the fault for not learning to be placed upon the teacher. To step in to teach the mathematics lesson will at all times waste the time of at least half the class for there is too great a diversity in student levels of maths here. Students could work on their workbooks in class and seek help as they came to difficult problems, but then this would leave the more difficult work, statistics, to be undertaken at home. Some students have already completed all six workbooks.

Disappointment: It is not possible for the teacher to know an individual’s mathematics and statistics work just from what transpires in a group situation. For each three hours in class there is a minimum of six hours out of class reviewing students work, changing the next step in the teaching to modify the mis-perceptions, to elaborate more fully, to follow students’ questions. That which passes between teacher and student by way of homework, by way of practice examples, is the major mechanism for communicating what is understood and what is not by each individual. Students must use that mechanism if I am to help them and if I am to know their work.

I believe that all the students in these classes are of sufficient ability to gain the skills that are required of them in mathematics. I will work with all the students in these classes. However, to do this students must start the process, mastering by themselves as many of the topics as possible and in so doing identifying those aspects for which they need help. The cry that 'I need to have it all explained, I know nothing' simply signifies that the student has assumed no responsibility for their learning. This would suggest that the student had failed to meet the first objective of the mathematics component ie that students are responsible for their own learning. (Writings, Spring, 1994)

In week 4, with obstinacy and disappointment vented in the teachers diary, the intended teaching was put aside and the students' concerns discussed. They were not concerned about their mathematics as the head teacher had portrayed. It was statistics that they did not understand!

Students at this time had completed the first three statistical units. The first two homeworks had been discussed as in the previous section and returned to them. The homeworks, as returned, had lots of comments to extend the work. They were not graded. The units thus far were:

- You need no introduction to statistics!
- Seashells on the seashore (Generating the statistical process); and,
- Comparing the seashells on the seashore.

In week three, the students were given the task of *comparing* the seashells on the seashore (refer video clip *Comparison*). A diary entry noted students:

> … were given the task of making the comparison of the shells on the beach before and after the construction of the breakwater. One group completed the exercise, the remaining two were to learn from their 'failures'. Failing to communicate with the other team working on the problem. When the time came to compare results there was no comparison was possible. What has been learned from the exercise has as yet to be communicated to me. (Writings, Spring, 1994)

Students perceived that they did not understand statistics, yet the ideas that were the substance of the units were ideas that they had generated, albeit cued by judiciously selected prompts and exercises. To tackle this issue of 'I do not understand' the students were asked as a group to re-construct what we had done in class. 'What did we do the first week?'… 'What did we do last week?' The students had no problems reconstructing the statistical ideas that they had generated. It was as if these ideas were too easily understood. There had to be something more profound that they did not understand. It seemed that there was an expectation that the easy ideas were not what were to be learned, there had to be something they did not understand.

Students at this stage were grappling with how to identify and write about the statistical ideas they encountered. As the teacher, I perceived that it was an inability to write about these ideas that was actually the problem not the understanding of the ideas. In a video clip, *Sampling,* collected the following week, students returned to this issue of understanding versus the ability to write. In this clip, two students are seen to concur with idea that writing, not understanding is the problematic issue.

5.4.8 Student challenges to learning

Confidence in their own ability.

Throughout the course, it was apparent that students were not confident in their own ability to generate good statistical thought. For many students this was evident in their use of textbooks. They had observed and participated in activities, they could articulate the statistical ideas so generated, but they would not or could not commit their own observations to paper. The search of the text often included ideas not covered in class.

A second instance of this lack of confidence was exemplified in the fourth week by a student who maintained it was not possible to 'think about statistics' before being able to 'do mathematics'. This student had been captured on video-tape in the previous class using insightful statistical argument to persuade his group to carry out a task in a particular manner. He had at that time convincingly argued, that if they took a rough quarter-sample of a much larger sample of shells that the results would be comparable to having worked with the original sample from a much larger beach population of shells.

Remembering rather than recreating.

For others, being required to reflect focussed them upon their ability to remember, rather than the ability to recreate ideas from the recollection of the initial stimulus. Mick sought help for his poor memory. In class he could only remember a few ideas to jot down. When asked to reconstruct the important issues given the stimulus question, Mick's ideas (with my penning) filled three blackboards! Not only this, his unleashing of ideas was in an order that was logical. The ideas flowed. The ability to communicate is of paramount importance. Yet the exercise was of itself not sufficient to help Mick unleash the ideas in the subsequent unit. More work needed to be done.

Interference by psychological reactions

For other students, the homework evidenced emotional reactions to the work, but conveyed little as to the student's ability to reconstruct or whether she does indeed comprehend at the point of instruction. Gwen's homework did not address what happened in the activity. She wrote:

> I find that when I am being taught I can comprehend but it just doesn't register afterwards thus getting me confused and upset because I know I can do them but then I draw a blank. Come time to mark I realise where I went wrong and feel totally stupid as I was always told. But I'll stick to it because one day something must sink in.

Gwen's statement is perhaps more appropriately made in relation to the mathematics component of the course (references to 'do them' and 'come time to mark') than it is to the statistics component, for which a comment was requested. Here, I felt there was considerable work to be undertaken.

Despite many years teaching experience (supposedly good teaching), the impact of the teacher's communication on students was exposed using this approach as it had not been in many years of traditional teaching. Regurgitation or the massaging of my words or of others' texts were no longer taken as being indicative of what the students understood. I needed to see evidence of students combining their ideas in ways that were their own, rather than mine, in order for me to accept that their writing reflected their understanding and observations.

Students needed to be able to distinguish between their discomfort in the process or learning and their ability to understand and utilise the statistical and mathematical ideas.

5.4.9 Outcomes

Over the ten week period, students were required to participate in activities in order to discover, through observation and class discussion, fundamental statistical concepts. During this time, students were required to submit five minor pieces of 'reflection' homework of the kind already discussed. Full marks were gained for submission of the work. Late work was accepted up until the tenth week. A minimalist approach with these submissions was taken by only a few, and the development over the homeworks was often considerable (as will be illustrated in a case study in Section 5.7).

In the first three weeks, the deliberate failure to provide students with a set of notes that could be regurgitated was a source of frustration and anger for students, particularly when it was apparent that their homework fell short of the exemplary homeworks selected for placement on the noticeboard. Students maintained that 'they did not understand'. In week four, the challenge of this 'I do not understand' statement began. The statement was challenged, whenever it was uttered, through statements such as 'What did we do in class?'…'What happened when we did that?'…'What did you observe?'…'Tell me then, what it is that you think you do not understand?' and if there was a lack of understanding then the ideas were identified and clarified (refer, video clip *Comparison*). After the third of these homeworks (Week 4), there was a detectable change in atmosphere: homework was viewed as a source of feedback rather than assessment; and, students submitted optional pieces of work in an effort to improve the communication of their ideas in their writing. The finishing standard for most was appropriate for commencing 100 level university students.

In breaking away from a more conventional approach of students having to learn technical material without having to grasp the concepts underlying this material, the subject was a qualified success. The qualification is that, in the main, the students did not fully appreciate their success. Many did not realise the complexity of the material they had grasped. It was as if there *HAD* to be something they could not understand; these 'simple' ideas could not be what they were to understand! Whilst student performance was exceptionally pleasing, most did not perceive that the real gains in a tertiary preparation course were learning how to learn and, in this instance, learning to deal with uncertainty. Student ratings of various aspects of the subject deteriorated (as will be discussed in Section 5.5.6, after the third implementation).

In the preparation stages of the course I had reminded myself 'that this thesis was not about metacognition'. At the end of this implementation I was aware that students needed to understand the process of learning and what it entailed. For the third implementation, the course was renamed *Learning Mathematics and Statistics*. What the teacher-researcher was attempting to do in changing the name was to legitimise in the students' minds their experiences and frustration in learning to construct their own concepts and meanings and to trust their own thinking.

5.4.10 Opportunities for collegial feedback: Implementation 2

Collegial feedback was possible after this implementation through the presentation of a paper titled *(Re)constructing Knowledge in Statistics through Reflection, Articulation and Language* (Porter, Griffiths, & Hedberg, 1994c). At this stage the feedback did not induce changes to the next implementation.

5.5 Case study 3: Learning Mathematics and Statistics

In the second implementation, the improvement in mastery of the content and statistical thinking, as demonstrated through the students' discussion, activity in class, and through their assessment tasks was a teacher's delight. (A comparison of students' understanding of the disciplines and attitudes towards the subject over the three implementations is provided in Section 5.5.6).

However, there were problems in the second implementation. From a teacher's perspective, students appeared dismissive of their learning. It was as if the ideas were too easy, when students had anticipated them to be difficult. They did not realise that the challenge had moved from mastery of difficult mathematical and statistical concepts to facing the challenges of learning in a new way. It was possible to identify many of the sources of concern for students. They included, amongst others:

- establishing relevance;
- dealing with feelings of uncertainty;
- expressing mathematics and statistics in writing; and,
- dealing with a number of feelings associated with learning.

They had overcome many of the learning challenges that had confronted them. They had been involved in learning how to learn but seemed unaware of this. Whilst there was evidence of an improvement in understanding, students perceived the subject more negatively than for the previous session. (This evidence is discussed fully in Section 5.5.3)

There were choices to be made in how to deal with the students' concerns. It would have been possible to change the mathematics and statistics curricula and/or teaching practices. However, the ideas generated by students and the connections made between them were the best the teacher had ever witnessed. Whilst the teacher sought to eliminate some sources of student concern, others remained, with the emphasis on providing an environment in which students were better equipped to deal with the learning challenges. The approach to and rationale for dealing with two of these challenges, establishing relevance and dealing with uncertainty, are dealt with more fully.

Relevance

Some challenges, such as helping to establishing the relevance of the materials was a challenge taken up by the teacher-researcher. Establishing the relevance of the discipline was interpreted as an important first step in engaging students in the learning of the discipline. Relevance was, by the third implementation, perceived as having three components. These were identified as:

- *Discipline relevance.* Students with a clear perception of their future discipline of study often wanted examples from say, psychology, in order to establish relevance to one's future career.
- *Ownership*. Students often generated their own data and/or the means for dealing with data and this held some relevance for them.
- *Authenticity*. This conveys a sense of the material or activity being important or personally meaningful. Clayden et al (1994) wrote about the term *authentic activity*. The usage of the expression coincided with the sense of personal meaning that I attempted to engender in students by establishing that statistical thinking was important in everyday life. Clayden et al (1994) discusses the sense of authenticity as it applied to the following example.

> Sam is five years old. He enjoys school and he particularly enjoys maths. Observed during a series of maths activities, he was seen to work extremely hard at the tasks in the commercial scheme used in his classroom. This work predominantly involved drawing objects to make up sets and then colouring the objects in the set. When asked what the work was all about, he said 'colouring'. He called his maths work-book his 'colouring book'. (Clayden, Desforges,Mills & Rawson, 1994, p.163).

The discussion centred on young children and what was important, that is, authentic to them. In this instance what provided authenticity was not the mathematics but the activity of colouring.

> …it does seem consistently to be the case that young children focus on the working practices of their activities in making direct interpretations of them. In their efforts to make sense of their classroom experience, the working practices are much more salient than abstract ideas. Colouring, discussion, and managing reporting-back are important things to

> sort out in the eyes of these children. What they learn from classroom experience is how to do work, how to be neat, how to finish on time (or sometimes to how to spin work out) and how to tidy away. (Clayden, Desforges, Mills & Rawson, 1994, p.164).

To more firmly establish the relevance of statistics taught in the *Gateway* to the disciplines and to everyday life:

- students were introduced to the *Decisions through Data* videos. In the second implementation they were used in the last five weeks as both enrichment and revision. (For one student questioning the relevance of this subject to his intended future career, was a near weekly event). The videos helped establish relevance. In the third implementation they were introduced in the first week and better integrated throughout the subject.
- Teacher talk continued, making reference to a range of disciplines, with connections to the student's chosen disciplines, everyday life.
- Students continued to generate the ideas and data that were utilised in class.

Establishing relevance and ownership appears to be a much easier task than creating a sense of authenticity. Experiencing authenticity through a walk on the beach or a good novel is easier for most than experiencing authenticity through the statistics discipline. The difficulty in providing authenticity in statistics for students can be exemplified with comments made by two students completing the *Statistical Literacy for Arts* program. Even though it was useful and they could see its value, they were simply not 'big fans of statistics':

> In school, statistics was about maths and formulations; I was not at all keen to do it over again at uni. This unit, however, shed new light in the application of statistics to the real world which has yielded more interest to me than just mathematical concepts… The experience was OK, statistics will always be statistics no matter in what form, which for me means it will always contain MATHS…Aaarghhh! (student, Statistical Literacy for Arts).
>
> I'm not a big fan of statistics, but I guess ARTS101 has helped my research and analysis skills remarkably and helped me in my other subjects. I thought the minor assignments were a drag at the time but they helped me think about my final task better. It gave me a chance to think about my angle of research…(student, Statistical Literacy for Arts).

Helping students discover authenticity through the learning of statistics, remained, even after many implementations, one of the teacher's most difficult tasks. The approach adopted in this subject was to create awareness of new ways of thinking, particularly statistical thinking.

Uncertainty

In the learning environment for the various implementations, *uncertainty* was a source of anxiety, or discomfort. However, the issue of dealing with uncertainty is interesting in the context of statistics. Statistics is a discipline that is non deterministic, many of the answers that students obtain are probabilistic. Students (and professionals), must be able to deal with answers which have attached a degree of uncertainty. Indeed they must be able to respond in a variety of situations, such as real life problem solving to the uncertainty of the situation, ascertaining what is relevant and what is not. Students need to be aware that the choices they make often lead to varying outcomes. The pedagogical practice, of using activity based learning, where students were to experience, observe and make sense of the situation, was a source of uncertainty. By the third week when students had to write-up and submit their observations, they wanted the safety of a set of notes to regurgitate. They wanted certainty.

The teaching practices could have become more prescriptive: giving more lectures, providing notes, providing more precise statements about homework; providing grades on weekly homework. The choice made was to structure the environment such that students were better able to recognize that they were learning to deal with uncertainty and the other challenges they confronted. So, rather than remove many of the challenges that the students faced in learning, in the third implementation of the statistical and mathematical literacy subject students were taught, explicitly, using the rubric of *learning to learn.*

The change to a focus on *learning to learn*, was one of emphasis, particularly in teacher talk but also through exercises which were more *explicitly* directed at students becoming aware of how it is they come to learn. In the third implementation there were several issues to deal with:

- how to prepare students to deal with uncertainty;
- how to develop their writing;
- how to help students distinguish between understanding and being able to write;
- how to help students distinguish between understanding ideas and understanding mathematical language;

- how to have students' focus on all major concepts rather than the easy or hard ones;
- how to develop students' ability to listen;
- how to orientate to students to creating and re-creating ideas rather than remembering;
- how to create relevance and particularly authenticity in the work; and,
- how to identify the students' silent cries for help.

The remainder of this case study:

- examines the techniques used to establish, for students, an explicit focus on *learning to learn*;
- describes the nature of the data collected;
- provides a comparison of the three implementations;
- identifies issues which could be examined through a re-examination of existing data or through subsequent implementations; and
- describes the teacher's insights into teaching and learning gained over the three implementations.

5.5.1 Techniques for providing the focus on learning how to learn

The framework for focusing on *learning to learn* was established through a variety of means:

- teacher talk, to *set the scene* and raise the issues which confronted previous students;
- exercises, to focus students on learning how to learn;
- creating instances to reflect upon current learning;
- assessment for learning, rather than for grading students;
- class discussion; and through
- providing a commentary on the experience of learning as various exercises were undertaken.

The techniques for drawing the focus to the process of learning were at times subtle, at other times obvious, took seconds or were sometimes part of a specific exercise. At other times, it was indirect, as a by-product of a statistical exercise. Sometimes it focussed on how student knowledge itself was constructed and at other times on the emotive nature of the process. However, whatever techniques were utilised it was important for students to be aware that the learning process was considered as important as learning the material itself, as learning was to continue when the teaching instruction stopped. Not all the techniques were new to this implementation. However, the focus on learning to learn was more *explicit* than in the second implementation.

Setting the scene

In week one, the students were asked the simple question, 'What is the name of the subject that you are about to undertake?' The unanimous response from all three classes of students was *Mathematics and Statistics* rather than *Learning Mathematics and Statistics*. This perception was rectified through an initial lecture explaining the approach to teaching and learning that was to be adopted. Usually, a couple of students will indicate at the end of class that they 'feel better or more relaxed' after this explanation, their mathematics anxiety having diminished a little, although this is not necessarily the case for all.

During the introduction to the subject, students were informed of the approach to learning that had been adopted as portrayed in Exhibit 5.21.

Exhibit 5.21 Approach to learning

My teaching is premised on the assumption that each of you as learners will take what I have to say and put it together to form different constructions of knowledge, different perspectives on that knowledge. This premise is a basic tenet of constructivist theory and it contrasts with the notion that each of you as learners are passive vessels into which we pour knowledge, store it until it is again released in its uncontaminated form. The communication of what it is you understand is crucial if I am to help you develop appropriate representations of that knowledge. To aid that communication I have lots of class discussion and minor writing assignments… Approaching teaching from the perspective of constructivist theory, I require you as students to participate in activities and to make sense out of the ideas that those activities engender. In our work we will attach labels to ideas and repeat them week in week out so that you come to understand the language and ideas of statistics. I will not give you a lecture, sets of notes, readings…I want you to think about what you observe in class. I want you to piece together a jigsaw puzzle of ideas, to develop an understanding of what statistics entails. What are the most central ideas, what ideas are of lesser importance? You will need to learn to describe what it is we do and what we have observed, talk about it and write about it. (Introductory talk, implementation 3)[6].

The students were also introduced to the issues that confronted students in the previous two implementations. These 'teacher-identified' issues involved students:

- trusting their own ability to think about and do mathematical and statistical work;
- ascertaining the relevance of the material to their chosen discipline;
- ascertaining the personal relevance of the material to their lives;
- modifying expectations about the discipline when it was not as expected;
- dealing with uncertainty as characterised by 'tell me what it is I have to know, give me a rule or a formulae to follow';
- being able to reflect upon what they had experienced in an appropriate manner, that is to be able to focus on the most important concepts or ideas rather than focussing on the easy concepts or the hard concepts, or upon the specific example rather than the generalisation or vice versa;
- in being able to recreate ideas rather than simply remembering them; and,
- in being able to write about statistical ideas.

Exercises for focussing on learning

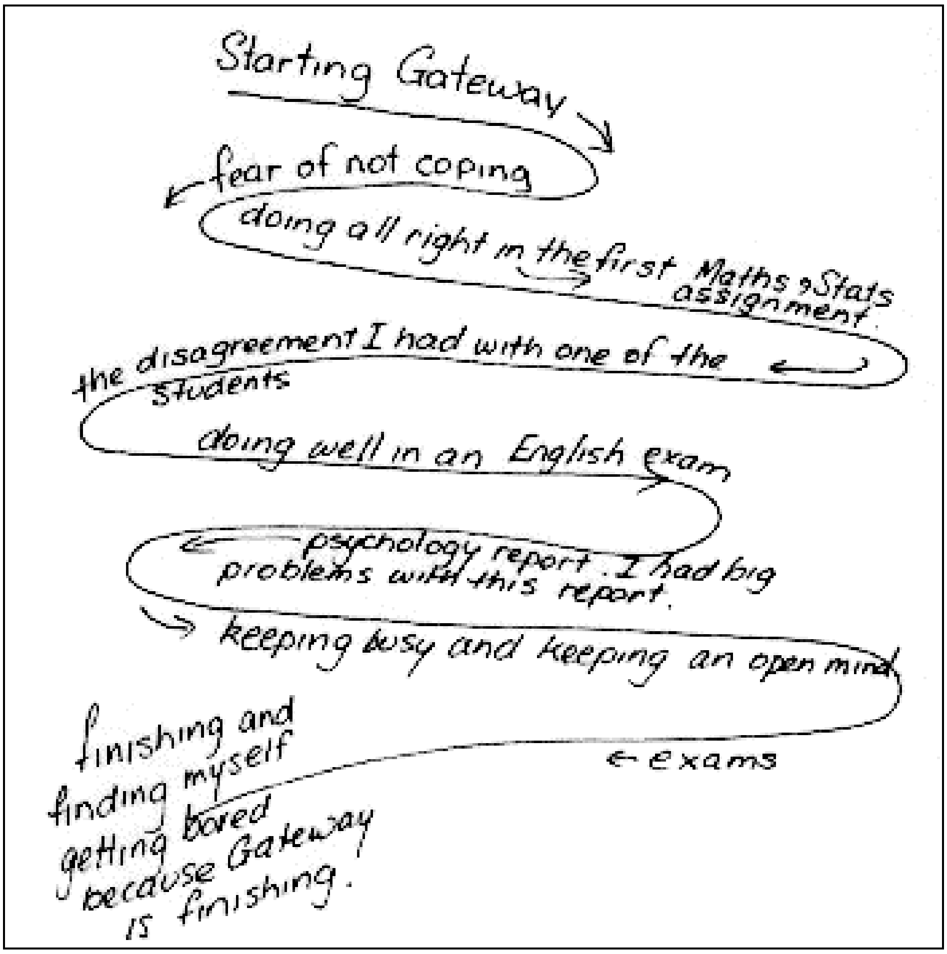
- The snake. Also in the first class, students were asked to draw a winding snake or stream (Pope & Denicolo, 1993) and on it to chart the various incidents that previously influenced their learning, culminating in the experience that led to them

[6] This has been described more poetically by Kellaway (1922)
The teacher with his watering can
Of formulae and words,
Of theories, theorems, epigrams,
Hypotheses and surds,
With cunning, wisdom, wit and stealth
Pretends to rear the perfect man
Exactly like himself

But what he's really looking for
Are funnels stuck in ears
Down which he may hopefully pour
(to boost his ego, calm his fears)
his facts and figures, data, dates,
the knowledge he anticipates
will cancel all arrears.
"An Image from Montaigne" (1922)

applying to enter the *Gateway* program. This exercise was to focus them upon those incidents that had helped or hindered them in previous studies, see for example Figure 5.5. At the end of the subject, this was repeated with students being asked to focus on the journey through the subject, see for example Exhibit 5.22.

Exhibit 5.22 A student's journey through the Gateway program



- Personal challenges. In the first class, students were asked to anticipate 'What are the personal challenges that you think you will need to face and conquer in order to learn (1) your Mathematics?; (2) your Statistics? and (3) the material in the other units in Gateway?' Repeated at the end of the subject, this exercise provided material that was then used at the commencement of the subsequent implementation.

- Sharing student histories. For the tertiary preparation groups anxiety appears to be the highest during the third and fourth week of session, when the first piece of assessment has been returned and another piece is due. Students, feeling uncertain, want to know what they have to learn and what they are to give back to you. Using the stories gathered from previous students the class was alerted to the commonality of their experiences. [In later implementations a selection of the end of session snake drawings was shown to students in Week 4. These were extremely effective in illustrating the paths that students follow the range of emotion they may encounter and the fact that, with perseverance, they emerge successful at the end of the course.]

Reflecting on one's current learning

- *During tests.* When the tension could be felt in the middle of a mathematics test, students were asked to stop, close their eyes, and note their anxiety level. They were reminded that the tests help them assess how much they have learned. They were asked to think about *why they were anxious*.

- *Mid subject.* Similarly, midway through the subject at the start of class the following instructions were given:

  > Close your eyes, breathe deeply for a few moments...now think about the challenges you have to face in learning your mathematics and statistics...*pause*...What are they?...*pause*...Take a few minutes to write them down…*pause for a few minutes*…Would anyone care to share what their challenges and the strategies you are using to overcome them?'

  Problems such as my partner is complaining that I spend too much time studying are identified. Strategies such as, I promised my partner that I would take a week holiday when I finished, are shared between students. Teacher talk is used to introduce the strategies shared by students in earlier implementations and to draw more solutions from the students. Other strategies to deal with the difficulty in understanding the statistics included for example finding a textbook such as 'Statistics without tears' or for understanding mathematics, doing some examples everyday.

- *Morning Tea.* A 'morning tea'/ social gathering was scheduled in the fourth week, between two one and a half hour mathematics and statistics teaching sessions to focus on learning issues. One issue arising concerned the kinds of *support* provided by partners. Students were alerted to the subtle (or not so subtle) differences between supportive statements. For example, 'Does you partner say to you *I won't feel less of you if you decide to drop out* or do they say *don't worry about the kids, I'll take them out for a play*?

Assessment for Learning

What students learn is often influenced by the nature of the assessment tasks. The assessment tasks in this subject were integral to the learning process. The accompanying activity sheets asked students to reflect not only upon the activities conducted in class but the experience of doing them, for example, 'what did you learn about yourself whilst working in a group situation?' Full marks were given (as in the previous implementation) for four pieces of statistics homework, if submitted. The returned homeworks provided students with feedback as to the quality of their work, clarified or extended ideas, showed how writing could be developed to an appropriate quality illustrating the depth and breadth of ideas that were required for their major assessment work. Case study 4, in section 5.6, follows the development of statistical thinking and writing as a student completes the four statistical homeworks and the major statistics assignment.

Class Discussion

'How are you going?'…'What was the purpose of the work last week?' These questions were frequently asked at the start of class (and on meeting students out of class). These questions engage students in discussing and sharing their experience of learning. A recent example of a discussion, after a peer assessment session involved a student asking the question, 'Will you be giving us written feedback on our assignments?' Not satisfied with peer assessment, following a teacher generated set of criteria, observing other assignments which could provide ideas as to both content and layout, many students wanted to be told what was required by the teacher. This question was re-directed to illustrate an introductory week's discussion. The discussion was about the need for many students to be told precisely what to learn rather than to learn to select for themselves what is important in a situation, for the latter may invoke feelings of uncertainty and lack of trust in oneself. In the ensuing discussion some cited their awareness of the requirement that as part of the homework and peer assessment they were to identify for themselves what they believed was appropriate and to learn from the situation. Others become aware that they had self censored in their writing and learning 'I didn't know you would want graphs or that we could put in our own data'. They had attempted to give the teacher what they thought the teacher wanted rather than respond to their own sense of appropriateness of material.

Providing a learning commentary during class exercises

- Evidence of thinking statistically: Several of the class exercises provide the teacher with an opportunity to challenge students perceptions of themselves as mathematical or statistical thinkers. The exercise *Measuring Variability*, involves students working through a worksheet developing ideas of location and spread, using both given criteria and self developed criteria to assess which they consider to be a good measure. When the answers are obtained, for example, a zero for the sum of the deviation scores, or a large number for the sums of squares or an incomprehensible variance, students think they are wrong or stupid because they do not understand. Drawing attention to their thinking, indicating how it is that they are right and how they perceived the need for a better measure such as the standard deviation which they could understand, can be used to affirm the students as statistical thinkers.

- Teacher talk: Teacher talk can help. For this same exercise on measuring variability I could share my experience of thinking there was something wrong with me when I felt more comfortable in interpreting a logistic regression than a discriminant anlayis. My own experience was of saying to the professor-expert 'now why do you use this technique instead of that' only to be told what I already understood 'well the logistic regression is easier to understand'. That teachers can be waylaid by the same thoughts, the same self doubts as the novices can be instructive.

In a classroom with the teaching based on activity based learning their is a fluidity in teaching that is not planned, but anticipated, particularly after previous implementations. Planned activities and responding to the opportunities as they presented allowed the focus on *learning to learn*, through this process of learning mathematics and statistics, to be made more explicit to students than in previous implementations.

5.5.2 The early conclusions

Before a detailed comparison of the three implementations is presented it is useful to note that at the end of this implementation *no* major changes were anticipated for subsequent implementations. Such were the outcomes.

At the completion of the third implementation other issues and questions arose and minor modifications made to address issues such as *cultural inclusiveness.* As opportunities arose in the process of looking at learning issues, data collected from students were used to examine other issues such as the *usefulness of feedback* given to students and t*he impact of other subjects and events outside of the Gateway* that affected learning.

In presenting the comparison of the first three implementations, outcome data of two subsequent implementations are also included. There was a curiosity as to whether or not the results were maintained beyond the phase of collecting the primary data. The fourth and subsequent implementations allowed the process of reflection upon the primary data to continue throughout the writing phase. At times it also allowed an examination of specific questions of interest.

5.5.3 Comparing Outcomes

A major concern for the discipline expert, when teaching is addressed to other than the discipline content, is whether or not the learning of the discipline is maintained or enhanced. The response of *I loved the subject, learned about myself but didn't learn or understand any of the mathematics and statistics* was not acceptable. The three implementations may be characterised as:

(1) baseline;
(2) curriculum refinement; and
(3) learning to learn.

To examine whether or not the *learning to learn* implementation, would result in comparable assessment outcomes to that *curriculum refinement* and *baseline* implementations, a comparison of the three implementations was carried out. Some data were also available for two subsequent implementations. These are included to see whether or not the results were *maintain*ed in later implementations.

Data were available on several aspects of the implementations to enable a comparison of performance in the subjects. Data included:

- assessment marks;
- rates of assignment submission; and
- completion rates.

To provide a quantitative comparison across implementations of understanding, rather than a non-comparative description of the ideas that students had generated throughout the activities and homework, a comparison has been made of one examination question asking students to map the major statistical concepts.

A quantitative and qualitative comparison on the students' experiences of the subject, across the three implementations was also possible. At the end of the fourteen week period, students were re-aquainted with the comments made in the opening lecture as to the teacher's perceptions of what students needed to contend with in order to complete the course. Through the provision of their personal accounts of the learning journey they had completed they were then asked to confirm or deny these ideas. Specifically they were asked:

> What were the personal challenges that you had to face and conquer in order to learn (1) your mathematics?; (2) your statistics? and, (3) the material in the other units in *Gateway*?

The data from this exercise were used to provide a student interpretation of their experience of learning statistics. (The term *student interpretation* is used here, although it is, of course, compiled by the teacher-researcher). Students also completed an anonymous formal course evaluation, a series of rating scales (refer Table 5.8) conducted by the Centre for Staff Development, in the absence of the lecturer of the course.

Outcomes: Assessment

Any comparisons over the three implementations need to be qualified. A strict comparison of performance cannot be made as the assessment requirements altered slightly. The exam was made more difficult in the second and third implementations through the addition of an extra question. Whilst the number of assessable pieces remained constant, the composition and weighting of the assessment changed in implementation 3 (refer, table 5.4).

Table 5.4 Assessment and weightings

Assessment	Implemetation 1 & 2 Baseline and Curriculum	Implementation 3 learning to learn
Participation - class	20	10
Participation-learning task		10
Statistics homeworks	10*	10
Mathematics skills	10	10
Major assignment	30	30
Final examination	30	30

* For the first and second implementation there were 5 statistics homeworks, but for third *implementation* there were four homeworks

Various changes are evident over the three implementations as is evident in Table 5.5. These changes include:

- *Marks*. It is evident that performance in assignments and overall marks improved greatly in the curriculum focussed implementation when compared to the baseline (implementation 1), despite students' poorer pre-test scores in mathematics. This was maintained when the orientation of the course shifted to include an explicit learning focus.

- *Examination performance*. The decline in exam performance as a consequence of the extra question, added after the first implementation, is noticeable.

- *Submission of assignments*. One of the gratifying changes over the three implementations is with respect to the rate of submission of assignments. This has improved. In the third implementation, completing students submitted 100 per cent of assignments. The submission rate was not only maintained in the fourth implementation, students were submitting write-ups of more exercises than required!

- *Completion rates*. One of the disappointments in the third implementation was the higher dropout rate. It was possible that students who failed to complete were the mathematically weaker students. A two-way analysis of variance of the mathematics pre-test scores found no significant differences between dropouts and those completing the program ($F_{(1,193)}$=0.99 p=.3212), but did find a significant difference between the implementations ($F_{(2,193)}$=17.7 p=.0001). The students in the first implementation had a higher mean pre-test score than either of the two following implementations. As a teacher of the only non-humanities subject in the *Gateway* there was a constant followup, wherever possible, to ascertain whether this strand was the cause. The decline in completion rate in the learning-focussed course was explained in terms of outside influences eg eviction, illness interrupting students' study and family difficulties.

- *The variation in marks.* In exams, assignments and overall marks, there is also evidence of a reduction in variability. It had been noticed throughout the final implementation, there were fewer high fliers, fewer poor performances and a shift upwards in the bulk of marks.

Table 5.5 Subject assessment outcomes

		Implementations			
		1 Baseline	**2** Curricula	**3** Learning	**4** Maintain
Number commencing		64	79	61	60
Number completing		47	62	37	48
Mathematics	mean	82.4	64.0	62.8	Not
Pre-test	sd	14.7	20.3	16.2	given
Course Completion	%	73	78	61	80
Assignment submission rate	%	76	95	100	107.5
Assignment	mean	7.4	8.8	10.7	
mark /15	sd	3.8	2.9	1.8	
Statistics part	mean	10.8	8.6	9.2	
of exam mark /15	sd	5.1	2.6	2.4	
Overall mark	mean	71.3	78	81.5	77.6
/100	sd	13.6	10.1	6.3	8.2

It is not possible to isolate the causes of differences in performance. The improvement in overall marks from the second implementation reflects the increased assignment submission rate and occurred despite the more difficult exam and lower mathematics pre-test marks. The increase in submission rate may have been due to the simple change in implementation 2 of starting the process of reflecting (and writing) in class at the completion of the activities in the first few lessons.

Outcomes: Understanding

Whilst assessment marks are assumed to reflect understanding and mastery of statistical concepts they may also reflect students' ability to write or to echo the teacher. To examine the nature of student understanding of statistical concepts, the concept mapping examination question (refer Exhibit 5.6) was examined more fully. Students were to identify and link the major concepts. The responses to this question have been examined so as to ascertain the frequency that:

1) variability was positioned as a central concept; and
2) variability was included as a major concept, but not as the central one.

For each implementation, the proportion of students who provided an answer to the examination question, and the proportion of students who presented variability as a central concept or a major concept are presented in Table 5.6. These data reveal that:

- while the baseline implementation had one less question, only in the second and third implementations did 100 percent of students complete the question;
- only 12.5 percent of students in the baseline portrayed variability as the central statistical concept, rising to 51.3 per cent in the learning to learn implementation;
- the percentage of students who included variability as a concept, whether central or major, rose from 47.5 percent in the baseline to 78.3 percent in the learning to learn implementation; and
- the percentage of students indicating variability to be a central or major concept was higher for the learning to learn group than for the curriculum refinement implementation.

Table 5.6 Percentage of students identifying variability as the major statistical concept

Implementation	Question	Exam	Central	Major	Total
	n	n	%	%	%
(1)Baseline	40	46	12.5	35	47.5
(2)Curriculum	62	62	40	24.2	64.2
(3)Learning	37	37	51.3	27	78.3

Students' answers may also be examined to reveal the proportion of students including each of the major concepts used to identify the statistical process. The proportional representations of the other major concepts, ethics, types of questions, design, measurement, sampling, data analysis and conclusion, are in Table 5.7.

Table 5.7 Percentage of students identifying the concepts in the statistical process

Implementation	n	Ethics	Question	Design	Measure	Sample	Analyse	Conclude
		%	%	%	%	%	%	%
(1)Baseline	40	0.0	30.0	5.0	75.0	100	82.5	52.5
(2)Curriculum	62	16.1	46.8	71.2	83.9	91.9	87.1	53.2
(3)Learning	37	10.8	56.8	78.4	70.3	91.9	89.2	70.2

Again it is interesting to compare the proportion of students providing the concept of analysis, 82.5 to 89.2 percent compared to the proportions including variability, 47.5 to 78.3 percent (as summarised in table 5.6). Although representing and summarising means was an integral part of all exercises, there was an explicit focus on variability in each exercise. Despite this, analysis is represented as a major concept more often than variability. One explanation could be that initial conceptions regarding the nature of statistics is one of analysis, calculation and representation of data. It appeared more effort was necessary to create an awareness of variability as a major statistical concept, providing the rationale, t*he why of statistics* than it was to create an awareness of design, sampling and analysis.

Outcomes: Student satisfaction

In the second implementation, the improvement in performance and understanding was accompanied by a decrease in student satisfaction with the subject according to the independent evaluation (refer Table 5.7 and Table 5.8). In every respect, students found the subject less satisfactory than in the first implementation. The challenge in the third implementation was to maintain the improvement in marks/grades whilst increasing students' satisfaction with the subject. As the data, all ratings are higher for the third implementation than for either of the first two implementations (refer Table 5.8).

Table 5.8 Indices of student satisfaction

	Implementation				
	1	2	3	4	5
The lecturer sequences material in the unit very illogically (1) to very logically (5)	3.9	3.6	4.3	4.3	4.6
The quantity of material has been far too little (1) to far too much (5) (3 is the preferred response)	4.1	4.3	3.4	3.1	3.8
The lecturer stimulates me to think about the subject never (1) to always (5)	3.7	3.7	4.1	4.1	4.4
The lecturer has demonstrated understanding this subject very poorly (1) to very well(5)	4.3	4.3	4.7	4.6	4.8
I have felt enthusiatic about attending lectures in this subject never(1) to always (5)	3.8	3.3	4.4	3.8	4.5
The subject material has usually been very uninteresting (1) to very interesting (5)	3.6	3.3	4.1	4.6	4.5
The lecturer has presented the material very unclearly (1) to very clearly (5)	3.7	3.2	4.1	4.2.	4.5
This lecturers interest in assisting me to learn is non existent (1) to very high(5)	4.0	3.9	4.5	4.5	4.8

An alternative representation of these data involves the examination of the percentage of students who respond more positively to the subject (refer Table 5.9). The students perceptions, whilst showing some variation in the fourth implementation, are consistently higher in the fifth implementation, with all students finding the subject interesting! This possibly reflects the teachers greater comfort in teaching with the activity based approach and to smaller refinements to the subject. Students percieved little difference in the amount of material covered in the first two implementations. Students in the third implementation did have one less statistics homework to submit. The percentages of students considering the courses to be in the range, slightly too little work given to slightly too much, for the three implementations was 93, 94 and 97 percent of students respectively. There was little perceived difference in the amount of work given.

Table 5.9 Percentage of students responding to the two highest rating categories in the formal evaluation.

	Implementation				
	1	2	3	4	5*
The lecturer sequences material in the unit very illogically (1) to very logically (5)	76	47	97	86	85
The lecturer stimulates me to think about the subject never (1) to always (5)	70	51	76	75	80
The lecturer has demonstrated understanding this subject very poorly (1) to very well(5)	86	85	91	93	90
I have felt enthusiastic about attending lectures in this subject never(1) to always (5)	63	45	88	64	90
The subject material has usually been very uninteresting (1) to very interesting (5)	55	40	82	68	100
The lecturer has presented the material very unclearly (1) to very clearly (5)	60	38	79	84	85
This lecturers interest in assisting me to learn is non existent (1) to very high(5)	70	66	88	89	90

* Beyond the fifth implementation, the rating scales used by the independent evaluations changed from using a five point scale to a six point scale. Thus comparability beyond implementation five was not possible.

The three implementations were sequential and did not follow an experimental research paradigm with subjects being randomly allocated to different conditions. The teacher-as-researcher paradigm used the reflective process and associated evidence to direct changes to teaching. As such, changes are not interpreted as causative of outcomes. The evidence suggests that it is viable to take time in the mathematics and statistics classroom to focus upon issues related to learning, as this does not necessarily detract from the outcomes of learning the discipline. This is important as far too often in teaching there appears to be too little time to cover the content let alone explore other issues such as *learning to learn*.

For the subject, which was part of a tertiary preparation program, to be considered a success, it was seen as necessary for students to be more inclined to undertake further study rather than less inclined. Mastery of content, alone, was considered inadequate. One of the advantages of including a focus on the learning process was that there was an improvement in student's attitudes toward the learning experience.

Outcomes: Learning challenges

It was apparent that as the subject has changed, so too did the issues that the teacher and students confronted in class. In this section there is an examination of the issues which confronted students and how these changed from one implementation to the next. For example, in the first and second implementations the teacher identified the issue of authenticity. In the third implementation, the students did not identify it as an issue. Perhaps this was because, in this implementation, some eighty-two percent of students found the subject interesting. It is the teacher's interpretation that the learning framework provided this element of authenticity and hence it was not salient for these students. It could have also been that some students experienced a change in their perceptions of the world as a consequence of the subject, and gained their authenticity through developing and refining their ability to think statistically.

Relevance was not raised by students as an issue throughout the course patter and discussion during the third implementation. In the final questioning of students regarding the challenges that had confronted them, a few comments on relevance were provided:

> If the relevance of the subject could be demonstrated to the other subjects, eg Psychology, I think it would have connected more.

> ...the first five weeks where we had to explore. I didn't know what! I didn't see any relevance.

It is possible that the video clips from the *Decisions through Data* series, more aptly interspersed in the early phases of the third course, did, as intended, help to satisfy the need for the relevance of Statistics.

In previous courses students had sought counsel because of their inability to remember (rather than students perceiving the need to recreate). There is no apparent reason for the absence of this problem in this implementation. Certainly students did not refer to it as problematic in their statistics component as they had done in previous implementations.

Other challenges that were identified by the teacher as needing to be addressed by students were endorsed by students' descriptions of the challenges they faced. The characterisation of those challenges has in several instances led to finer categorisation of the problems experienced.

The category of trusting one's ability to think statistically was retained but differentiated from, more general statements pertaining to confidence and self-esteem. Student comments related to trusting their own ability to think statistically were as follows:

> I believed statistics was totally beyond my understanding.
>
> Lack of confidence in my academic ability yet determination to overcome it...I have been taught from an early age...to accept the fact that I will continue to make stupid errors in calculations.
>
> Stats... was not my forte...I found it hard to digest therefore I was not as open minded as in other subjects.
>
> My preconceptions concerning the level of difficulty alienated me in that I never realised until half way through that I had some understanding of what I was learning.
>
> I was apprehensive about my mathematical capacity. I was very much afraid of it (maths)...I had to psyche myself up to thinking that I was going to get over the obstacle of fear, and I did. My attitude started changing when I realised that maths & stats was like a puzzle and it was a challenge.

These comments were closely associated with, but differentiated from more general trust and confidence in oneself as the following comments reveal:

> Could I actually do what was required of me?
>
> The more concepts were introduced...the more I became unsure of whether I could do it. It was all new.
>
> I lack self confidence and hold a negative attitude towards myself which was also a challenge to me.
>
> I need self esteem and confidence building.
>
> The problem I face is that I can't start a piece of work until I know the idea I have is correct...I don't believe I can do it myself.
>
> Even now I still have trouble coming to terms with my negative thoughts...believing in my own ability to achieve will continue to be a challenge throughout my studies despite my success in the Gateway program.

The need to modify expectations was differentiated into expectations about Statistics and expectations about themselves as learners. Students in earlier classes thought of

Statistics as information, particularly information in graphs and tables. Few were aware of the processes involved in producing statistical information as exemplified by the following:

> I thought Statistics was just reading graphs.
>
> I had a preconceived idea of what it was going to be about, mainly graphs and how to read them.
>
> At first I thought Stats was only things like 5 billion people live in China.
>
> I could not make sense of the concepts. Then a revelation...that Statistics is not definitive. It felt like the world had been lifted off my shoulders.

Expectations about themselves as learners also needed to change at times:

> Accepting that the material was something that I didn't know...
>
> Accepting that I didn't know everything and that it was OK that I didn't know and that learning is the process of knowing.
>
> Prejudice (belief that Statistics is unapproachable) is a strong feeling that can't be set aside so easily.
>
> I have realised that a person can go about obtaining an answer depending on interpretation. I realised that I could have my interpretation and someone else could have their interpretation of the question and still arrive at the same answer. I have also thought that my interpretation and approach were wrong. I thought that by being told how to do it I couldn't go wrong. However I also discovered that by giving my ideas a chance I wasn't incorrect.

Many discussions amongst the teachers of the *Gateway*, over all strands (literature, critical thinking, psychology, history...), suggests that learning to deal with uncertainty is a challenge of gigantic proportions. Students readily concurred with this in the following comments:

> Not knowing what was expected ... she made me learn and discover things for myself ... the negative feelings might have disappeared had she given us the formula and answers ... but then we wouldn't have learned anything.
>
> I was uncertain about my own understanding.
>
> Not knowing whether I was on the right wavelength.
>
> At the beginning self doubt as to whether I was capable of this.
>
> I need to get a specific answer to a specific question.
>
> I would still love more structured teaching...as I felt I was grasping for straws and was never quite sure if I had grasped the concept...the fun ways statistics was taught left the pressure of learning behind while the class moved toward actual statistical concepts.
>
> There are definitely feelings of uncertainty; can I do this?

> ... and the uncertainty of 'is this what I want?'
>
> I think I was lost at times because I always wanted to be spoon-fed but now I realise that I had to tackle statistics the way I knew how and keep persevering with it. ... I'm so glad I was pushed to think and was not told the things I wanted to be told. I don't think I have conquered the uncertainty but I think I've become aware of it and I am not as concerned about it any more.

The ability to reflect, that is to identify and organise major concepts, was portrayed as a difficulty by five students (a difficulty which was not evident through their work). In this implementation, students were far better at addressing the major concepts in an appropriate manner than were students in earlier courses. Perhaps, discussion of this early in the course averted the difficulty; perhaps the course has matured; perhaps students were more aware of the learning process. The challenges identified were:

> Shaping concepts into a meaningful, coherent array of knowledge.
>
> Instead of looking at the small parts of the statistics units, I was looking at the whole, and that was daunting.
>
> Each individual concept in the statistical process appears very hard to understand. (a top student).
>
> I had problems...putting all the concepts into some logical order.
>
> In all the units in *Gateway* I faced a dreadful sense of inability to understand the general concepts.

The challenged identified as, *the ability to write*, was differentiated from a more general statement regarding the use of language, although it is likely that the ability to write is also closely associated with the grasp one has on the statistical or even English language. For many, the challenge of writing was indicated:

> When asked to write down the subject matter it became difficult.
>
> The ability to write in Statistics is a major hurdle.
>
> Found it hard to write about statistics.
>
> I found it hard to get the presentation right in statistics...I was not focussed on the actual subject.
>
> ... learning about statistics was OK but writing about statistics stayed demanding.
>
> My language does not communicate what I want. Understanding how to write a statistical report.
>
> ... my writing ability was not up to scratch. ... I was confused as to how much information was to be included into the report, and the category the information was to be under.
>
> I didn't really have any trouble understanding the concepts involved, but all through the course I've had trouble writing it down.

> Writing assignments in the correct language and style.
>
> My inability to express myself was evident when writing statistical reports and essays.

For others, the language itself, either the statistical language or English language was seen as problematic:

> I had trouble grasping the statistical concepts due to the labels. I understood what variability, bias and other concepts (as words in english) are but not as mathematical/statistical concepts.
>
> Statistics was a completely new area to me and I amazed myself at the terminology I began to understand.
>
> The belief that spelling is SO important.
>
> Coming to terms with the english language, and conquering my fear.
>
> Stats allows me (within reason) to physically handle the material in a way that I can assimilate it to learning (the overflow has forced me to rectify my writing).
>
> I am from a non english speaking background and sometimes I do not understand.
>
> I had difficulty in interrelating statistical information with the work we had done in class. I wrote down definitions of statistical words etc and didn't apply them to what we had done in class.

One student identified *alienation* as an issue. However it may be more apt to ask students who dropped out of the course what the challenges were that they so briefly confronted. For this student the sense of alienation was as follows:

> In terms of alienation ... in class discussions ... fellow students have ... years of experience ... I feel rather devoid, useless to the conversation.

Other challenges surfaced. Not perceived by the teacher was the challenge of *overcoming the gap since previous formal education*:

> The lack of formal education for many years.
>
> I had been out of tertiary education for 20 years.
>
> The problem of starting after years of no schooling and the feeling of stupidness compared to recent HSC students.
>
> ... dealing with the art of learning after so many years.
>
> ... my age, the gap from the school years.
>
> Fear was my biggest obstacle ... I had not studied for 12 years.

Students were concerned about *new material*. It is interesting that in a course that is to provide the background to University studies, to provide a level of literacy, that failure to have background knowledge, was of concern. It was expected that students would not know about Statistics in the early phases of the course. However the newness of material was a concern for students. The challenges as they expressed them were:

> In the early stages the concepts were mostly alien to me.
>
> I felt like I was starting from scratch when learning statistics.
>
> Statistics was a completely new area to me.
>
> I struggled with statistics. This may be because I have never touched on the topic and it takes time for me.
>
> Statistics! I knew nothing about...but had heard that it was quite unnapproachable.
> It is hard starting with brand new concepts.

Other commitments such as *family and work* resulted in, *time pressures and tiredness*. This was not unexpected it had been the focus of discussions throughout the course, but it had not been returned to the students as a challenge they would have to face. Students faced the challenge of:

> ... prioritising and organising the workload.
>
> Work and family ... bearing on the amount of time.
>
> I need to schedule and re-schedule time.
>
> Time constraints.
>
> I found it very demanding fitting in the problems of my family life.
>
> There were so many hurdles, weeks of pressure and trying to juggle every facet of my life.
>
> Sleepy during videos.
>
> After a busy day at work ... I was constantly tired.
>
> I couldn't find the time.
>
> I have a problem doing assignments because I am a housewife and I had to have time away from my children.

Also the focus of discussions in class was the problem perceived, by the teacher, of the *perfectionist* for whom work would never be good enough. Students also had a perception of a grading system that would give high marks, as in school. Good performance seemed more akin to full marks than credit grades. Whilst not commented on to a large degree, there was often evidence of this problem in class. Here one student states her situation:

> I realised I have very high expectations and that I am a perfectionist. This was a difficult realisation for me, but one that I had to accept so that I could accept that the reason for the mental block was my own attitude.

In addition to the challenges of *gaining confidence and self-esteem*, there were other challenges in the affective domain to conquer. As the students commented:

> ... my weakness in personality.
>
> ... overconfidence in exams (implied leading to silly errors).
>
> ... emotional overload.
>
> ... the creativity of coming up with ideas.
>
> The biggest challenge I had to face and conquer in order to learn was my self-abusive thinking ... I continued to remind myself that I have every right to be here. I had to stand my ground because for the first time in my life I found something that I can do.
>
> I felt an overwhelming fear of failure and inability to overcome what was a mental block. I felt enormous personal pressure to succeed.
>
> I felt too shy to ask questions.
>
> Dealing with my frustration.

For others, isolated challenges were provided:

> I found statistics a challenge basically because at secondary school I found maths totally boring and hard.
>
> … concern about being accepted into University.
>
> Feelings of no support from internal family...actually made me more determined.

Apparent throughout the students' commentaries is the value of interaction between students and between student and staff in examining their concepts and emotions when learning statistics. The light-hearted patter that took place in courses was often cited as a source of personal or academic revelation. The following student comments illustrate this point.

> A breakthrough in my own mental blockage/ something which was holding me back. I wasn't getting it!! This came in the form of an unexpected question ... when we were asked to write down the problems we were facing, both personal and otherwise. The following day I noticed a change in myself. For me it was monumental as it heralded a totally different outlook and sense of confidence to achieve and understand.
>
> Knowing that everyone else was in the same situation eased my tension.
>
> My comment to the teacher that my family comes before Gateway was answered with 'that's easy to say but ... This made a significant change in my attitude and allowed me to free myself from some of my guilt.
>
> Talk. All the class have similar problems. I felt reassured, felt better.
>
> A revelation happened by chance on overhearing another conversation that Statistics is not definitive. It felt like the world had been lifted off my shoulders.

Outcomes: Conclusions

From the teacher's perspective, the material in the second implementation was sequenced more logically, the class discussion and assessment suggested that the students were coping with the material and understanding it better than students in the previous class. In this sense the course was extremely successful. (These perceptions were confirmed by analyses in the next section.) No changes had been made to the Statistics curriculum in the third implementation.

Preparation courses can only provide a minimal excursion into any discipline. A successful outcome therefore requires at the very least that students are satisfied with the learning experience and that students are open to these ideas in the future. One could further hope that students develop a propensity to continue learning the discipline, to study it through choice rather than compulsion. This might not be possible for the greater proportion of students.

The third implementation focussed on these latter objectives but without sacrificing the statistical learning objectives. For this group of students the focus on learning appeared to provide a source of authenticity. The strategy adopted as discussed was to focus on self as a learner of Statistics. Through this, it was thought that the learning experience would gain greater authenticity than in previous implementations. In contrast to the previous implementations, the relevance of Statistics did not appear to be a major concern of students.

One issue had not been satisfactorily resolved. Students were not provided with notes and readings in this course. This was intentional as many students are prepared to rote learn such material in the belief that they are not capable of understanding this work. Students in these classes had been required to think about what they observed in class and were to write about it. This was used to reinforce the notion that students can (and must) think statistically. Students do, however, require supplementary material. It is likely that computer simulations of similar exercises would be useful, although not yet available. Reading material, apart form other students work and teacher commentaries, introduced during the latter stages of the course is also likely to be useful in the process of refining ideas. However, what and when it should be introduced is still to be discerned.

At the most recent graduation, a student from an earlier intake of students addressed the graduating class. The student compared University studies with that of the *Gateway* program. They had in *Gateway*, the student described, experienced and worked through intense emotions, but now it was time to put this aside and to trust in themselves (or what they had done in *Gateway*). It was not possible to sustain this level of emotional intensity. It was, I interpreted, time to focus on the academic material before them.

5.5.4 Postscript: Cultural inclusiveness

At the end of the third implementation, but before presenting a paper (Porter, 1995) at another conference, I thought that the final revisions had been made. Indeed the materials were printed for the first class upon my return. However, that was not to be the case. Experiencing the culture of the aborigines of the Northern Territory of Australia allowed another focus to form. There were eight aboriginal students in the third implementation of the Statistical and Mathematical Literacy strand. All eight dropped out of the *Gateway* program.

The statistics component remained broadly unchanged, although a question formed as to the suitability of the activity-based approach for this group of students. The changes to subsequent implementations involved the replacement of the mathematics pre-test. Some, but not all, of the aboriginal students had extremely disrupted early schooling, and hence very poor mathematical skills. For many students, the pre-test was an effective means of initiating work. However, the possibility that it could cause students with weaker mathematics backgrounds to be early dropouts was too great for it to continue.

The pre-test was replaced with an exercise, *Identify the mathematics strategies*. Students were given the questions and answers to a series of mathematics problems. They were asked to form small groups and to verify or provide the strategies through which they could check the answers. In this manner the *learning* component was extended to include the discussion of cognitive strategies for solving mathematics problems. Students used different strategies and often strategies different to the teacher. This was useful for those students who thought that they were incorrect if they did not use the same strategies as the teacher.

Whether or not the change led to an improved learning environment for the aboriginal students was not ascertained. There were no aboriginal students in the following implementation. There had been successful completions by aboriginal students in the first two implementations, but it was not possible, after the event, to ascertain the proportions completing or dropping out. The practice of using the exercise, *Identify the mathematical strategies,* has continued.

There were other possible explanations for the drop outs amongst the aboriginal students, apart from overwhelming feelings of anxiety and uncertainty. They also had an alternative access program and did not need to complete the *Gateway* program in order to gain access to the University.

5.6 Case Study 4: Writing to learn

Asking students to produce write-ups of the activities (reflective homeworks) and providing full marks for any attempt (even if only the name of the student and title of the exercise were submitted) is not a common approach to grading. The reflective homeworks were part of an assessment package, leading to a major assignment and final exam, wherein quality and correctness were important. The approach to marking and assessment that developed throughout the experience of teaching *Gateway*, was one of marking for the students' *inclusion of appropriate issues, questions or concepts.* If a student addressed an issue, it could be clarified, refined, extended, corrected or questioned by the teacher. It was considered important for students to acknowledge issues and concepts as they arose even if they were portrayed in an incorrect manner. Recognition on the part of the student that there was a question, or decision or something to consider, was interpreted as the first step in learning. The parallel for this is evident in statistics, where experienced statisticians may interpret data differently, but they at least know what aspects of the data or process need to be examined in order to make an interpretation. The first step in learning was judged to be the identification of the important issues, and only then, the satisfactory resolution of the issues. Only in the final assessment was it considered essential to reasonably interpret issues.

The purpose of the homework was presented to students, as a mechanism through which they could develop their writing. It was to provide a means of communication about what they had learned, allowing the teacher to clarify any incorrect perceptions. It was not to be a piece of writing wherein they attempted to gain a high grade or avoid a fail grade, by glossing over what they did not understand.

This case study is presented in order:

- to illustrate the manner in which writing was utilised in the teaching and learning of statistics;
- to demonstrate how the minor homeworks were used by the lecturer to help improve writing and to clarify the statistical ideas encountered; and
- to illustrate the changing of student attitudes as they progress though the subject and through the *Gateway* program.

Initial connection with the student

This case study is based on one of the students from the second implementation. It was a student whose potential difficulties emerged early in the session. There was an instance of what I term a *connection*. I had a connection with this student in week 1, which made me become aware of who he was, his name, face, strengths and difficulties. Some students do not make such a connection. The connection between the teacher and student was over the manner of undertaking long division in mathematics. The student informed me that he usually got the answer, although he did not use the method illustrated in class. He demonstrated his method to me. My response was 'that is fine but if you want a quicker method you might like to try one of the alternative approaches we used in class.' I interpreted that he found this an unexpected response and further that in giving him permission to use his own method, I had helped open the door to his using alternative methods.

At this stage, I was aware of the various connections that I as teacher had with students, students with teacher, and students with work. These were significant moments in making me part of the teaching and learning classroom. It opened a door for me to work with students in a student-focussed manner. Somehow, through the connection, I see their humanity. However, that there was a connection did not necessarily mean that wondrous things would happen.

As detailed earlier, the students were basically asked to reflect upon the activities undertaken in class and submit the written-up reflections as homework. Where possible, they were to demonstrate the ideas using other examples. Students were given no instructions regarding page length or formatting the document. They were asked to explore their writing, submitting work that had met their own standards. The feedback would be used to clarify statistical ideas and suggest ways of improving the writing. It was the vehicle for their communication of ideas. The only further point of clarification was to provide an example of writing that conveyed meaning, (refer, Exhibit 5.5). After the return of the first homework, students were asked to compare commentaries. Each of the papers , good and not so good, had a commentary that could help the student improve their papers. The next step for improving the current paper was provided.

Student's progression in writing

Four of one student's reflective homeworks and final assignment together with the teacher's comments have been included to illustrate the change in writing for this student.

Homework 1. The first homework was received in the second week, after students completed the exercise, *The Statistical Process*. The student's sense of 'failure' expressed in this homework (refer Exhibit 5.23) was disconcerting. His sense that the group had not identified any statistical ideas, was an all too common response.

Exhibit 5.23 Homework 1

Student homework #1

The class exercise we did in class about the seashells brought home the idea of statistics and their concepts in a way that I hadn’t previously thought of. We had a short brainstorming session and came up with a group of questions that we thought needed to be asked and answered before we could continue. These questions were similar to the ones that are asked whenever any task is to be completed or formulated. Questions like how big is it, where it comes from, when, how long has it been there etc. many of these questions came up in the sea shells exercise.

I had an experience when we were moving something in to our unit. I needed questions answered before we could begin the process, how wide- deep long is it, how heavy is it, where was it kept etc. This experience led me back to the seashell experience. In our group discussion last week we came up with plenty of ideas and questions but I felt we failed to capture the ideas of statistics, our questions were very general and didn’t lead us to any statistical ideas.

The teacher's commentary at this stage was aimed at re-directing the student to the ideas the group engendered for they were indeed of a statistical nature. His negativity was accepted, but he was to justify his commentary in future. The teacher's commentary was as detailed in Exhibit 5.24.

Exhibit 5.24 Teacher's commentary on homework 1

In assessing this piece of work, I would want the student to provide their notion of what statistical ideas are for in this instance the ideas which we identified have been dismissed as too general and non statistical. In fact the questions that needed to be answered in order to undertake this study were many, detailed and possibly of such an extensive variety that as researchers there would be difficulty in conducting a study that was sufficiently comprehensive for any conclusions to be adopted as adequate. The types of questions that were asked are the sorts of questions that must be addressed in any study. If the real purpose of a study is not clearly established, the design, sampling and measurement undertaken in an appropriate manner then any reporting of results (which is what many people perceive as the statistics) will be inadequate.

Homework 2. The second homework (refer exhibit 5.25), even shorter in length, was prepared and submitted in week three, after students completed the exercise, *Making comparisons*. A task listed for homework, apart from the basic reflective exercises, was as follows:

> Over the next week collect from newspapers and magazines, examples of graphical presentations of data. Which of these graphs are easy to understand? Which might have been useful for displaying the seashell data? Why?

This same week, the student received his written feedback from his first homework. Along with the other students, he also received a verbal commentary in class that profiled the types of approaches taken to writing up the homeworks. At this stage, there was no improvement in the student's writing.

Exhibit 5.25 Homework 2

Student homework #2

I collected my examples of graphical presentations of data from the Sydney newspapers on the week end. they were collected from the sports, business and general interest sections.

They are nearly all simple examples of bar and line graphs. They all display the data correctly and in an easy to read and understand format. The best example I found was one showing the NRMA's profits after tax. It was clear, easy to read and displayed all the relevant data required.

I would have liked to use this example of a vertical bar graph to display the sea shell data as all the relevant information could have been easily displayed within the graph and presented in a manner which is easy to read and understand for anyone.

The Teacher's commentary was brief, in response to the brevity and quality of the student's homework:

> Try sketching your data in the same form as the various graphs you have collected. This will help deepen your analysis. What is it that we actually see in the graphs?

It was also suggested that by *actively doing* something, sketching his data, that he may develop more ideas. Exemplary homeworks such as that provided in Exhibit 5.6 and Exhibit 5.7 were also placed on the noticeboard, together with the commentaries for the student.

Homework 3. The third homework, on the unit *Sampling*, was submitted in the fifth week (refer Exhibit 5.26). For this homework, the teacher produced a general comment sheet (refer to the earlier Exhibit 5.4) detailing the issues that had arisen and also discussed the homework in class. This was supplemented by specific comments on each paper. In this homework the student included a reasonable representation of statistical ideas, which were extended and clarified in parts (see the footnotes and bracketed additions) throughout the text.

Exhibit 5.26 Homework 3

Homework #3
In class during the previous week we[7] conducted an exercise to dicover the mean age of the class. The exercise involved designing a study that *we* could conduct, deciding the measures used to conduct the study, true and accurate collection and collation of the results, discovering the best method to display The information by the way of graphs so that a conclusion can be drawn, displayed and proved. The exercise proved to be very useful and brought up some symbols that we previously hadn't seen as well as some new terms. The exercise conducted helped my understanding of the terms as I was able to use and apply them in a practical sense. Some of the terms that we came across and there meanings were, sample - this means the group or objects that you have taken the sample from, sampling is used in statistics. Instead of choosing to measure every object in your study you may choose to take and examine a smaller number. This is sampling. Population was the total number of people *(or observations)* used in our survey *(or study).* Variables - there were many variables in statistics. Some can be controlled, others such as weather etc can not, these must be taken into account in statistics. A statistic is a figure used to represent one or a number of things. It helps to display information. Mean, *(an example of a statistic)* this is the average of a series of scores. Population parameter *(we used the mean of a sample to estimate the population parameter ie the population mean)* refers to the particular group being studied. Estimation can be used when guessing a particular result, it can be used to help design the study. Sigma notation is the act of replacing numbers with symbols. There are many symbols used in sigma notation. Some of these and their meanings include $\overline{X}$ *(the mean)*, the end figure, the equals to of an equation, Xi, ΣXi (the total addition of a number of figures), n refers to the number of sums that are to be used. An example of this would be if four students in a class were taken as a sample, there ages were 10,17,22 and 15. Ignoring any variable such as their sex, you were asked to work out their mean or average, this would be *(the total of the ages)* 10+17+22+15=64, *(divided by the number of observations)* then 64÷4=16. The mean

[7] This and all subsequent occurrences of 'we' is circled and it suggested that the first person not be used in writing.

Homework #3
would be 16. Sigma notation for this would be *(Note check my summary for notation)* Sigma notation is often used in statistics. Its use is popular as it simplifies the writing of an equation which is particularly useful when the equation is large, involving many numbers. After beginning to gain an understanding of sigma notation I could identify and now use this when doing calculations in everyday life. In our class exercise *we* draw samples of people from the population and compared the mean[8] to that of the whole sample *(replace with population)* group or sample of the class. The exercise showed some interesting conclusions. (1) the conclusions that could be drawn from the exercise were, the larger the size of the research and sample the more accurate the result will be. If two people were drawn and a mean calculated this would not be as accurate a study as if two thousand were taken. Evidence of this could be drawn from the fact that the mean of the first four people I chose was 32.4. The overall class mean was 29.66. When researching try to survey or analyse as large a number as possible. (2) the other conclusion that I drew from the exercise was the larger number of scores we took the closer we got to the class mean. Everybody in the class was asked to take a sample size to measure, nobody came up with exactly the overall class mean. This can be used to prove that the only way to have gotten the class mean was to have measured all the scores.[9]

In this homework, the student demonstrated control over the statistical ideas and it was possible to comment on the writing. Hence the suggestion to avoid the use of the first person. At this University, undergraduate students undertaking writing tasks are for the most part discouraged from using the first person. Subjects which use qualitative techniques and theory, and which encourage or expect the use of first person writing, tend to be in the minority and are not usually on offer to students until later years of study. At this stage, no explanation was provided and this was probably was probably a mistake on the part the teacher.

The teacher's specific comments to the student reflect a dialogue that continued in class. Hence it was somewhat cryptic. Students' interpretations of these types of comments and the use of rhetorical questioning in feedback became a focus of the fourth implementation. The comments on homework 3 for this student are as detailed in Exhibit 5.27.

[8]An extension is inserted - of the samples

[9]An extension is inserted - That is, as the sample size increased the variability tended to decrease as the sample means converged on the true mean.

Exhibit 5.27 Commentary on Homework 3

There is room for improvement, but much better than your last two pieces.
Plotting the group's data could have helped.
The exercise was similar to research more generally. Most research is based on a single sample. This exercise showed that samples provide innaccurate results, but possibly good estimates, particularly if the sample is larger rather than smaller.
Writing needs work, but most of the ideas are included.

These comments differ from the first two sets. In this commentary a comment is made upon the quality of the work, 'much better than your last two pieces.' Commentaries for the most part were kept neutral, not providing reward or praise. The intent was to extend all work, to create an atmosphere where students felt *short changed* if they were not shown how to improve their work. It was hoped that students would develop an intrinsic sense of reward, namely satisfaction that they had submitted their best given their time, circumstance and ambitions. It is a system that has been retained, although at times softened for students who appeared to need the external reinforcement.

It is of course entirely plausible, in the absence of detailed instructions as to what was required in the written homeworks, that the student had simply modelled his work on the other students' exemplary homeworks. It is possible that with an initial set of detailed instructions that the students first homework may have been better. This is a viable explanation, although my perception was that initial instructions would not have led to a better piece of initial work. Whilst there was anxiety attached to the first pieces of homework, many students moved toward reliance on their own judgement and style, knowing that I would *always,* even on high quality work, provide critical comment. (An examination of the interpretation and reaction to feedback is undertaken in the next case study.)

Homework 4. This homework was shorter than the previous one and focused on statistical definitions. It was completed after the unit, *Measuring Variability*. This was essentially a make-up homework. The student missed one of the activity-based homeworks. This homework (refer Exhibit 5.28) was based on an activity that was a pencil-and-paper exercise. In the exercise, students work through different ways of developing and selecting measures of the centre of data and variation in data. The homework draws on a certain amount of written material that was provided.

Exhibit 5.28 Homework 4

Student homework #4

In class during the week and in previous weeks, the class has been looking at aspects of measurement and how they may affect, firstly our research and secondly the data analysis. When the study has been designed, the question to be asked has been decided, and the method of collecting the sample and the size of the sample has been collected the research, descriptions and conclusions can begin. One method used to assist with the analysis and the conslusion is to find the middle of the data. This is done with the means of mode, mean and median.

The mode is calculated by finding the most common score. The score that occurs the most. The mean is calculated by finding the total sum of all the scores and dividing by the number of scores in the data. The median is calculated by finding the middle score when all the data are arranged in order of magnitude. The mode median and the mean are all methods used to measure the centre of the data, but they may not always be accurate.[10] Unusual numbers may however upset these measurements although the mean is the easiest to use where appropriate.[11]

Unusually large numbers either too large or too small can effect our research and data analysis. These numbers do not fit in with the others and affect our results so a method must be found to take this into account and lower its impact on the research and therefore the analysis. A measure must be found to find a way to measure variability.[12]

This can be done by squaring the deviations. However this number may often be large and may be of little relevance. The best way to measure the spread is by finding out the standard deviation. This is an accurate method because 70% of all observations will fall within one standard deviation of the mean…95% of all observations will fall within 2 standard deviations of the mean. Almost all observations fall within 3 standard deviations of the mean, When the standard deviation is found then our research analysis and conclusion will become more accurate.

Apart from the within text comments, the teacher's specific commentary, for the student is as detailed in Exhibit 5.29.

Exhibit 5.29 Teacher's commentary on homework 4

You continue to improve the writing. Further elaboration is necessary on some points and some aspects of measurement are omitted. I think the next step in your writing is to write from the perspective of an objective observer. Give outcome, explain and exemplify.

When variables are measured, they are measured on some measurement scale. Scales are sometimes defined as nominal, ordinal, interval and ratio scales. A nominal scale is ----- and examples of this would be the measurement of species of shells ie we simply name the seashells.

See the comment sheet also.

[10]This sentence until the end of the paragraph can be teased out a little more.
[11]The mean is more susceptible to the influence of these outliers (unusual numbers) than the median.
[12]Clarification. When we measure variables we find variability in them. This may be purely inherent variability eg the plane flies different distances simply because it does. Variability may also be due to sampling or measurement. We need to measure variability in order to know whether the differences are unusual or not.

The major assignment. The four homeworks led to the major statistics assignment (refer Exhibit 5.30). Unlike the minor assignments, the major assignment was graded. This student's homework was given full marks though, in parts, it reveals flaws in the student thinking. The first issue considered when marking was the coverage of concepts and secondly the appropriateness of usage of those concepts. As for the reflective homeworks the comments provided were aimed at continuing the learning process, suggesting ways that the student could improve his work. Throughout the essay are footnotes. These provide the teacher's comments within the body of the report. These comments work on the ideas and language or request exemplification. At the end of the paper are suggestions as to extension of ideas, or in this case as to alternative perspectives.

Exhibit 5.30 Student's major statistics assignment

Student's major statistics homework

STATISTICS

Throughout the unit on statistics the group identified used and reflected upon a number of statistical concepts and ideas. Some of these included the idea's and concepts involved in designing a study, sampling, measurement, describing and analysing the data, variables and drawing conclusions. The ideas of this unit was to gain knowledge required to use these ideas and concepts. This was achieved through a number of practical exercises and research conducted by the group.

When designing any study there are certain questions which must be answered before the study can begin. We must first know what we really want to find from the study, like the group exercise on the seashells. The questions raised by the group explored the issues that researchers needed to address in order to answer the question set out by the research. The questions that were asked could be grouped into questions that could be related to -

- The nature of the real issue, what is the question really about
- Design and control issues, this includes such things as when the data should be collected
- Sampling, when questions regarding design have been answered the collection of the data may proceed. The questions here that are raised are where and how much data should be collected andhow should they be collected.
- Measurement Issues, some of the questions that can be raised here are how will we measure the data. These can be done in many forms such as size, weight, colour. These questions also gave rise to the notion of the different measurement scales. That is nominal scales for naming varieties, ordering scales for measuring any possible deterioration in the data, continuous scales and specific ratio scales where we can use an infinitely divisible scale to measure the length size or weight of data.
- Ethical Issues, who has commissioned the research and for what reason. This could affect all aspects of the research, the sampling recording and the conclusion.
- Decision Issues, how much difference must there be found in the data before it is deemed to be significant. Any inherent or expected variance must be taken into account when making this decision and drawing any conclusions.

This process of designing the study was the first one looked at by the group. Studies are designed for many different purposes. They are used to research and gain some knowledge into what is being researched and studies.

Student's major statistics homework

Through the classs exercise on sea shells which asked the group to design a study measuring the effects that a breakwater would have on sea shells the group was able to build it's knowledge and apply it practically into the study. Although some problems were experienced it was overall a success as the group on a whole was more familiar with what goes into designing a study.

For the exercise on the sea shells a random[13] survey could have been conducted. this would have involved selections of an unknown number of sea shells being collected from the same areas of the beach.[14] Failure to do this will distort or bias figures. Sampling is an exercise conducted during most forms of research. The practical exercise reinforced to the group it's importance and helped to build on the knowledge that the group already had.

When the sampling has been completed the measurement process then begins. The sample size or any data can be measured in many different ways, colour, size, weight, shape or as with the case with the seas shells, the sea shells condition. When measuring in any research the same method must be used right through the research.

Three methods that can be used in measurement are mean, mode and median. The mean or arithmetic average measures the average of the scores by dividing the total number of the data by the number of the scores. The mode is the most common value found in the data. The median is the middled score when the data is arranged in order of magnitude. The mean is more commonly used as it is easy to use when appropriate. The mean looks at all the data and reduces it to one number, making comparisons between the data easier to draw. Practical exercise conducted by the group found that the more[15] data gathered the closer you will get to the mean[16].

These concepts of sampling and measurement were reinforced by a practical exercise that was conducted by the group. In class samples of size 2,3,4,5, and 6 people were formed. These samples were drawn from the entire class or population of students. Generally the groups samples were chosen at random. When the samples were chosen the group began to measure the data. The one variable that could be measured in each sample was age, which could be represented by the symbol Xi. The mean age was then calculated for each sample. This could be described as . Each such sample mean age, is an estimate of the population mean, age. Sometimes the real parameter is known. But most times it is not known.

What was evident from this exercise was that most samples did not give the precise population mean. As samples got bigger the estimates tended to improve and got closer to the group mean.

Now that the study has been designed the sampling completed and measured it is time to describe the data. The most common method used in statistics is by using graphs. There are many different forms of graphs, pie charts, bar charts, line graphs (polygons) and frequency distributions are some of the graphs more commonly used. Graphs assist in analysing the data as they arrange the data in an 'order' that is easier to understand. An example of a graph is in figure 1. This graph shows the mean age plotted against the sample size and provides an excellent portrayal of how the larger the data sampling size the closer to the mean the group got. It is important to be able to read and understand graphs as they are the best method to display data and they are widely used. Examples of all different forms of graphs can be found in most newspapers on a daily basis.

When the design process has been completed conclusions can be drawn. This is where the question of

[13]random - this will be difficult here.

[14]This is, if we are interested in finding the proportion of each species. If we are interested in the change in numbers then this method would not be suitable.

[15]should read 'sample data'

[16] continue mean 'of the overall population. As samples increase the variability of the sample means tends to decrease'

Student's major statistics homework

ethics may come into practice in the wording of the conclusion and possibly how the data will be presented may vary depending on who commissioned the research. As discussed previously the more accurate the research process the more meaning can be given to any conclusions drawn.

The group started the statistics component with some knowledge of statistics but this knowledge was generally without great depth. Through the group exercises that looked at some of the specific concepts like sampling, measuring and displaying the data this knowledge was built and then applied practically to specific exercises. As statistics begin to play an ever more important role in not just academic life but everyday life this knowledge will prove invaluable.

The teacher's commentary offered praise for the piece of work. It had conveyed, in a reasonable manner, most of the major concepts the students had encountered. However the comments also indicated how the student could extend his thinking still further. Footnotes have been attached to indicate some extensions and modifications of the work that were placed throughout the essay. The concluding comments were as detailed in Exhibit 5.31.

Exhibit 5.31 Teacher's comments on the major piece of homework for Statistics

15/15. Once again the standard of your work is considerably higher - think back to the first piece you gave me. This is an excellent capture of what we have done - with only a couple of omissions:
1) the role of probability in making decisions and as encompassing our assumptions about the world.
2) you could tackle the nature of the statistical questions
- is there a difference between groups in mean age?
- is there a relationship between height and weight?
- describe the level of cholesterol in Australian men?

The next improvement is to be able to move beyond what we did in class and to exemplify on a 'hypothetical but realistic' project. In doing this, don't lose the level of elaboration of each concept you have just portrayed in this essay.
Now that you have used the statistics process as the connecting the concepts, can you redo it with VARIABILITY used to connect the concepts?

Throughout the duration of *Gateway*, this student was, meanwhile, keeping a journal as part of his assessment for the unit, *Writing and Speaking*. With the permission of the student, an essay based on extracts from this journal was given to me at the end of session. It provides an account of the student's learning experience in the subject *Learning Mathematics and Statistics*. Whilst the teacher was noting the change in his writing and understanding of statistical ideas, the student was also attending to the experience.

Attitudinal change

The student script typifies the experience for many students. Many have had appalling experiences of mathematics. One student indicated that as a young child she was told she was stupid practically every day in one mathematics class. Students' anxiety rises rapidly until they begin to complete their workbooks. Most who commence the workbooks will gradually improve. There are always a few who need a consultation to break through the problems they are experiencing. It doesn't always take a lot to achieve a breakthrough with students. Len draws on diary entries throughout the teaching session in order to describe his experience:

> At school I always hated Maths. I had no interest or enthusiasm for the subject and so as a result my marks were poor. I knew this was a problem that would have to be overcome if I was to succeed at gaining admission to university. So it was with some trepidation that I began the maths and statistics course of the *Gateway* Program.
>
> The first few weeks of the course did little to alter my thoughts and feelings. It was like school revisited. We received the course outlines and some practise workbooks. At school I had problems with long division, multiplication, algebra and statistics, and they were all major components of this course. For the first month we were given class tests. My marks for these were not to bad, averaging around seventy percent. The questions I'd answered incorrectly were the ones relating to long division, multiplication and algebra. It was a nightmare all over again.
>
> Anne our lecturer for maths and statistics was helpful, but still found her difficult to understand. I felt that in class she tended to explain things too quickly and I would always miss something. I rang Anne and made an appointment to see her to get some help.

Students who do not respond early to this pressure might not satisfactorily complete the mathematics component. Their success depends on their mathematical background. This student was fortunate to have a 70% average. Many students' marks were much lower. For those who can't *do mathematics*, the experience when they can is often exhilarating. These students are often extremely grateful and delighted with their accomplishments. Len continued:

> Anne was very helpful and together we sat down and took a look at how to overcome the problems we were having. I left the hour long meeting with Anne with a new understanding and outlook on maths. For the first time in my life I understood long division, multiplication and maths.
>
> The results I achieved with Anne in one hour were far superior to anything that I'd managed in my six years at high school. I left the meeting with a newly found confidence and enthusiasm. While statistics were still causing me some problems at least now I had started to make some progress.

Students such as this who suddenly find this to be their preferred discipline are a joy for the teacher! However, whilst students learn to cope, the discipline often is still not enjoyed. Whilst Len's experience was of his improvement in Mathematics, the teacher was witnessing the improvements in his weekly homeworks. As he explains:

> With this newfound confidence with maths I started to put in some extra work. Whenever I had a bit of spare time I would do a couple of pages of our practise exercise material. The better I became, the more I practised.
>
> I began to gain an understanding of statistics. The class exercises had a lot to do with this, if you can see how something can be used in a practical sense it definitely helps with your understanding. It was good to finally make some progress in this area.
>
> The pinnacle for me occurred in our last class test. It was our most difficult so far and was based on everything we'd done up until then. I scored ninety per cent and it made me feel as though all the work had been worthwhile. I am starting to enjoy maths and no longer consider it a nightmare.
>
> I intend to use this experience to help me with other subjects. I began maths and statistics knowing I would struggle and probably keep on struggling all the way through. I did not have much confidence in my ability to overcome difficulties because I'd never done so in the past. But through hard work and perseverance I have succeeded. If I can take this attitude into my other subjects in the Gateway Program I know I will succeed.

Not all students are like Len. At times, when marking assignments, there was a pervading sense that students never listened! Statements like 'please provide an introduction to the topic' are duplicated, time and time again. Whilst on the one hand there was often an improvement as demonstrated through Len's homeworks, one must wonder what sense students made of the feedback provided to them.

5.8 Case Study 5: Students' interpretations of feedback

Writing was a battleground for students. It was where they struggled to put their statistical ideas together in a coherent manner. The reflective homeworks, in addition to their use in clarifying and extending statistical ideas, were used to shape and improve students' writing. Therefore there was a focus on how students develop their writing. Not all students showed the same improvement as Len (although the majority showed improvement).

The questions arose, 'Do students understand what I am attempting to say with the feedback?' In the fourth implementation[17], after having received feedback on a reflective homework, students were asked:

- How did you interpret the feedback?
- Did the feedback make sense to you? and
- What was your reaction to the feedback?

When students received feedback on their assignments, they were asked to hold them so others could see them, and in this way they could see 'the red writing on all work'. So as to see that other students also had comments to extend and improve their work, they were asked to read the comments on someone else's papers. They were asked if there was anyone who missed out on commentaries to show how to improve their work. There were none. Students were told that all papers would be provided with comments on how their paper could have been improved. They were also told the absence of precise instructions as to the form of writing was because students could often 'exceeded my expectations' by exploring and developing their own style of writing.

Students' individual, written commentaries were part of a four tiered feedback system. In addition to individual comments, sample homeworks were placed on the noticeboard, an overall written commentary or perspective as to how the homework was completed was provided, as well as a verbal commentary provided in class. Judging by student comments, the multi-layers of feedback appeared necessary.

Comments on feedback: Week 3 of Session.

It was evident that for some the individualised, written feedback was insufficient:

> I didn't understand some of them until Anne clarified them. I feel that I've possibly learnt something. The assignment in now, I don't think will be much cop. I think I have a direction for improvement and hopefully will follow this properly

---

[17] The fourth implementation, except for the replacement of the mathematics pre-test with an exercise on strategies used in mathematics and the focus on assessment feedback, was essentially the same as the third implementation.

> Design: this comment isn't expanded enough for me to understand it. The last comment I understood only after you re-explained. Description and analysis of data. Decisions in the face of variability. This comment I am not sure about where to slot it in.
>
> I have trouble understanding the comments. It is a bit clearer after discussion but I feel more practice would help me most. Perhaps an in class activity would be beneficial so I can reflect on what I am doing right /wrong.

That everyone did not understand was anticipated. The possibility remains for further examining what sort of comment, say the rhetorical question, or the cryptic comment, or other feature is not understood. However, it their reactions to the homework which students need to deal with. In a situation where there was no grade, and comments made to extend the work of all, the reactions were dominated by negative expressions of disappointment and confusion. They were unsure.

> Anne I was really disappointed I must have attacked this from a different angle. I would be interested in seeing the comments on assignment 2, to see if I have got it. Yet, I realise you are looking further into it than I did. I should have started from the beginning and worked through it, answering every question, the design, sampling, measurement etc. I would like to see how I go with the exercise I handed in tonight, and then maybe I should have a talk to you.
>
> Did I ask the question correctly? Still unsure of the terms ethics etc. Unsure of the phrase of process. Need to match process to writing. Pose the question and later take note of how we ask statistical questions. Unsure what you meant by that.
>
> The comment on my assignment has confused me a little. I understand that I needed to add an introduction and conclusion. I don't think I understand the ethics, expertise etc. Maybe I didn't understand the question.
>
> I was confused until you explained it. I thought there wasn't an ethical question because of the type of statistics I used weren't ethical. I understand now that I should have covered the question more. Maybe I shouldn't use something so familiar because I 'assume' that people have background knowledge.
>
> Do I need to learn statistics or what? I didn't think I understood the question that well. I should have explained myself better. I do feel a little rattled. I have found three statistical books up stairs that I am going to borrow. It would be good to sit down and go through the papers. Just for an explanation. I will keep trying.

Feedback for some, showing how to extend their work, was like a reprimand:

> I apparently don't know what is required of me. I'm not doing it right. I am not picking up what I'm supposed to do. I don't seem to have a concept at all of Statistics.
>
> My reaction was that I didn't realise that I did so bad. I'm disappointed in myself that I did not do as well as I should have. I think I should have read it over and over again and rewrote it. The comments you gave me were great I'm glad you told me the areas I need improving but some of your comments I did not understand could you please explain them a bit more for me.
>
> Does no tick mean that it was aweful? How can I change my way of thinking - only practice? My writing relates to letters, reports etc as is evident from my homework. Would an appointment with you help please?

Students interpreted the feedback as indicating that they did not understand, for others their lack of understanding was confirmed.

> It took me a while to do this and I thought it was a pretty good attempt but I did realise I hadn't applied all the concepts. The feedback was constructive as I now realise I should have displayed the data to illustrate the work and should have been more specific and in more detail.
>
> My reaction to my paper was "well I knew I was doing it incorrectly in the first place - and this just confirmed it!" … I understand you much more Anne on the comments that you gave to us as a group rather than the paper that I was reading. The written comments are good too.

Some now recognised where their submissions were incomplete or that ideas needed to be expanded but they needed to justify with a reason:

> I can see from your comments on my assignment that my report on the seashell study was incomplete because I did not cover every aspect of the statistical process. I do feel I need to say that I could not possibly provide a display of data, explain the mode of measuring the data and finally draw any conclusions on the data because the exercise was based on a hypothetical scenario. I did not have any tangible data to work from. Without actually fabricating data, I could not complete the exercise and I am sure you would agree that to do this would be to provide me with absolutely meaningless information. I can see that I could have done a more thorough assignment if I had applied these ideas to anther subject.
>
> I realised that I hadn't gone into it enough. Your comments were what I expected, but I still don't know where to go from here. What I did was revision of the class exercise in a different form. I didn't address the question. This subject has me absolutely stuffed (excuse the french). I will never get it.
>
> I realise that a specific example may have helped explain the process a little better. But I was unclear that an example was needed, which other people did. I thought I explained the process. This would be the process that I would use and what is 'the process is embedded?' Was it not clear? Hard to follow? or are you saying that I know it? I think I have done a scientific point of view style in this piece.

Some students responses to the extension of ideas was 'I thought I included that', as follows:

> Thought I had discussed method. After comparing last weeks to this weeks I Think I may have developed further. Unsure what 'go on' precisely means. Did this comment actually mean ideas as to how I would have gone about my procedure?
>
> I expected a "not too good" result as when we experienced statistics the week after handing in homework I learnt that I had not addressed my assignment correctly. The comments I received were constructive and pinpointed the areas needing improvement. I did however feel that a couple of ideas listed were covered but I assume inadequately.

Some focussed on working out what the teacher wanted in writing, rather identifying an approach to writing that best explained the concepts they had encountered:

> I wasn't upset by your comments put on my paper. I was a little disappointed though as I did spend a lot of time trying to work out what was expected and how I was going to do it. I was half expecting I had done it wrong because I wasn't sure what you meant by essay, lab

> report, whether to do it in an essay form like I did, or go under headings in point form like you suggested; I also didn't know whether you wanted us to do diagrams or not. But I'll know next time.

Others simply saw the potential for the feedback to expand and clarify their statistical ideas, rather than interpreting the comments to mean that it was wrong. Part of the work of the teacher has been to re-iterate that, virtually all work at a pre-tertiary and undergraduate level can be improved. My job, as I explain to them, as teacher is to extend what students have done and they should feel cheated if I don't!

> Glad for the feedback. I felt like I was working blind. Hopefully this weeks will be better. I think that I have a long way to go! I have included lots of variability in this next paper, and the sampling as well. Showed me where I needed to expand and to see where the different processes do fit in. Made things a bit clearer to me.

> I really like my feedback, it really gave me some good ideas for next time. I feel more confident now about my statistics work. I don't feel like I am confident about statistics, but I do feel I have more to work on. The table (graph) was great. I didn't even think of using a graph, but I will now. I actually discovered for myself about the cooking container heating up - I just hope mum hasn't yet! (Just joking)

> This week I seem to grasp the process of statistical studies more than at the time I wrote the above assignment. I'm interested in my lecturer's guidance throughout this part of Gateway. I felt more confident after the paper plane assignment. Look forward to Anne's further comments.

> Mine is OK. I agree with comments- how do I do it? I must use the right terms. It made me think how to. Encouraged.

It is extremely disconcerting for students to have their homework treated in this manner. In the third and fourth weeks, when their first pieces of work, in this and other subjects, are returned, their anxiety and discomfort peaks. Sometimes, for some students such as Len there is a perceived need to provide praise when there is an improvement. Where possible, the aim was to have students want the feedback to improve for the sake of learning, rather than attaining a particular grade. Certainly none of the students wanted to receive a fail grade, and many would have if their initial work were graded.

It was not possible to prospectively identify, from their writing, students who would drop out from those who remained. The comments of two students who later withdrew from the program can be examined. The writing from the first of these was from a student with a non English speaking background. Her writing suggested that she would have difficulties, although they were not necessarily insurmountable:

> I find in my report the idea in there a few mistakes (written and preparation). Conclusions is not very strong (poor) but by practice more I can become better in writing the ideas down and connecting them together. The other thing I have to practice more on the process which the report was based. To be able to separate and identify them 100% clearly from each other (discontinued study)

The second of the students who withdrew gave no real indication that they were to withdraw. The uncertainty 'Yes? No?' embedded in the quote did not suggest overwhelming difficulties for the student.

> I understand your comments. I must think to express or put in order. I need more order in the mind to show understanding. I think I have the basic concept. Yes? No?

Comments on feedback: End of session.

At the end of session, the concept of feedback as a mechanism to help students improve their learning (of all kinds) and their writing was evident. In response to the prompt, *what helped or hindered you to learn to write?*, one student's comment encapsulated those of most other students. The outside influence of other subjects, reading other students' assignments, feedback itself, the method of assessment (reflective homeworks) and the hands on activities were all commented upon by students. For this student all were seen as helping to develop her writing:

> Doing the psychology report and the weekly statistics homework really helped my writing for statistics. The feedback on assignments and reading other peoples' assignments helped the major assignment. The hands on exercises in statistics were very useful. As much as I hate to admit it the weekly statistics homework really helped pull everything together and made it quite clear what I did get and what I didn't get.

Whilst the teacher had focussed on feedback and the discussion in class, students had identified the need to clarify ideas, organise concepts, master the jargon and participate in the classroom statistics activities.

> Actually doing activities in class - active participation. Concepts become much clearer.
>
> I am good at writing and expressing things verbally but not in maths. It still helps to write sentences but the knowledge needs to be there because you can't waffle.
>
> The weekly statistics homework, helped me write specifically for Statistics. The comments received help me put the Statistics jargon into perspective.
>
> Writing in statistics is difficult if you don't remember the words to use. I find that I never seemed to go onto the next step.
>
> Writing in statistics was awkward. The terminology was foreign. I would have preferred grasping the new words first before we tried to attack the papers. My essay writing in general has a lot to be desired - so I have learned.

> Writing for Statistics was difficult. I struggled with ideas and concepts. Again, asking and looking at the other examples helped. When writing I would often struggle, have blocks and walk away from it. It was very frustrating.

Many found the feedback useful and were no longer negative, and this was as hoped:

> With the lecturers' help and guidance I have improved in my writing style, although it has been a difficult time, doing things sometimes twice or even three times. It has been disheartening at times, coming to terms with it, thinking that I really was not intelligent enough and all negative thoughts. I have overcome that now I believe with persistence and hard work I will succeed.

> In statistics the comments on my reports were good because then I could see what I needed to work on and what was included. In this way each one I handed in I learnt a bit more.

For the teacher-researcher the focus on what happened within the classroom was intense. *What was I doing as teacher?* And, *what were they doing as students?* This almost shut out the potential influence on learning from other subjects and life more generally. The teaching and learning of statistics was undertaken in a context of students completing other subjects and indeed, leading hectic lives. The impact of all the units within *Gateway* and how they interrelate may be explored by examining the turning points in the course. This was an aspect that could also be examined from data collected during the fourth implementation.

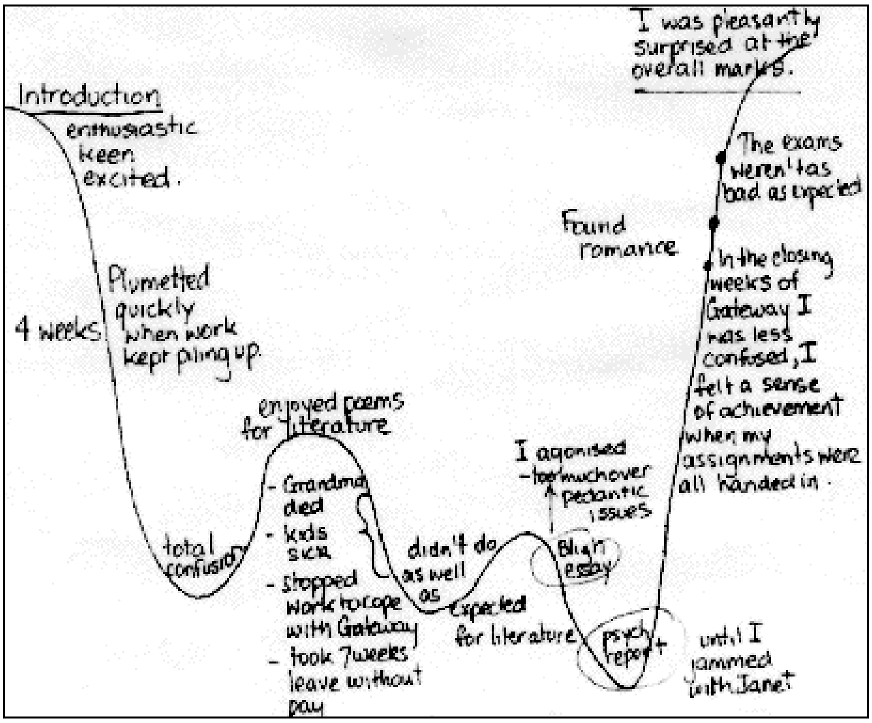
5.8 Case study 6: Turning points in *Gateway*

At the end of the fourth implementation of *Gateway*, students had responded to the cue:

> Draw a landscape where the hills and valleys are to represent your ups and downs in the *Gateway* course. On it, detail what the experiences were and how you felt and acted as a consequence.

The metaphor of the snake was modified as the turning points could not always be interpreted as having a positive or negative influence upon learning. The metaphor provided by the hills and valleys (and plateaus) was more suited to this task. One such learning landscape is presented in Figure 5.32.

Exhibit 5.32 Landscape of positive and negative influences on learning.



In order to understand the nature of the learning journey for the students the landscape were coded and analysed. The steps involved:

1. coding the landscapes according to the categories: start, incline, decline, positive and negative turning points, peek, lowest trough, plateau and end.

2. the start was coded to reflect the starting attitude toward undertaking Gateway. Four categories were needed positive, negative, mixed or not indicated. Positive affect (or emotion) was characterised by sentiments such as overjoyed, excited, interested, confident and ready to learn. Negative affect was characterised by sentiments such as dubious, terrified, daunted, scared, apprehensive and unable to move. Those classified as mixed expressed both a positive and negative sentiment such as *new job, interested in course and attentive to others and lecturers but very confused about objectives*.

3. the initial movement, incline or decline was examined in relation to the students' starting state.

4. the reasons or sentiments associated with the decline were characterised.

5. the reasons or sentiments associated with the incline were characterised.

6. the peeks were characterised.

7. the distinctly lowest troughs were characterised.

The landscape in Exhibit 5.32 was typical of many students whose starting experience was positive. In most instances, this state was followed by a decline in positive affect. This is consistent with the teacher's perception that the class level of anxiety was greatest around week 4. Captured in Table 5.10 is the startling and converse feature that, when students started apprehensively, in most instances they experienced an early improvement in affect.

Table 5.10 Changes in affect from those experienced at the start of *Gateway*

Start affect	Incline		Decline		Plateau		Total
	n	(%)	n	(%)	n	(%)	n
Positive	2	10.5	15	78.9	2	10.5	19
Negative	10	71.4	3	21.4	1	7.1	14
Mixed	2	66.6	1	33.3	0	0	3
Not apparent	3	42.8	4	57.1	0	0	7
	17		23		3		43

Multiple reasons were sometimes associated with the initial decline. The dominant reason was students' realisation of the workload (n=11), questioning the time available (n=2) and beginning to have doubts about the ability to study or complete the work (n=5) and the possible impact on family and life (n=2). The content was also cause for concern for students: mathematics (n=2), statistics (n=2), literature (n=2) and writing (n=1). Already for one student, one of life's events had intervened - her dog had died at the start of this new endeavour and she experienced grief and loneliness.

Those who experienced an increase in positive affect, after the start did so because they experienced the sense of 'can do' (n=12). They could understand, maths was easier than first thought, the library work was easy, they obtained their first high distinction. They experiences a fighting chance, a grasping new work (literature, mathematics language and psychology). They handed in the statistics homework. The incline for some appeared to be because, *they were still there despite* (n=5), still questioning how *Gateway* would fit into life, fear of presenting, uncertainty, workload and literature tutorial!

For those who faced an initial decline in affect, the reasons for the subsequent turn toward more positive emotions were far more varied than for those who initially experienced an increase positive affect. Emotional states improved when good grades or comments were returned (n=7). It improved when students got to know others and their journeys (n=2), *slowly I crawled out inspired with tales of woe from previous Gateway students,* and when lecturers took time to talk about the difficulties experienced by students (n=1). The turning point was also described in terms of the week, *in the 5th week I was more enthusiastic and optimistic*. Their emotions were more positive as they completed and understood the work (n=7), were able to understand/enjoy mathematics and statistics (n=5) and other subjects. At other times they resolved issues, *bargaining with the husband and children* (n=1), quitting employment (n=1), organising themselves, and simply picking up.

Distinctive low troughs were experienced. As evident throughout the many Gateway programs, *the lake of despair was after 4 weeks of Gateway. Ten weeks to go. So much to do I think I will die.* Often the students had to confront one least favoured subject or activity (n=11), *writing and speaking tutorial,* for another *STRESS, the Psychology report. What does Anscombe know anyway. I haven't even got all the papers. Will I make it into uni*. The dip in students' affect was associated with a number of factors. These included personal problems (n=3), the death of family members (n=1), disappointment with grades in mathematics and statistics (n=2) or in other subjects (n=2), being able to write, '*I could understand in my head, but couldn't put stats on paper*' (n=1) and other reasons (n=4).

The peaks of experience were often associated with a number of factors:

> Maths is easy. Got a pay rise and became a supervisor at work. Getting organised and I know when everything is due - not that it helps- I still do everything at the last minute - Best thing I realised - "ITS UP TO ME.

On most occasions, it had to do with obtaining good grades (n=9) or the experience of having learned, *The best experience in Maths was on the 2nd last week, Tuesday morning when I finally learnt the Pearson's correlation! What a wonderful feeling!!!* Sometimes it was associated with having completed work *or having fun, making friends*. Sometimes it was just a *good feeling about myself* or of being organised.

The students charted their inclines and declines in affect, their peaks and troughs. These were the students who completed *Gateway*. The interplay of completing a number of subjects was evident. There was some anecdotal evidence that most students who dropped out of the Gateway program did so for personal reasons such as death or illness. However there was also evidence of students surviving the *Gateway* experience despite such trauma. In this exercise the number of instances of negative affect attributed to life outside *Gateway* was somewhat less than the teacher expected.

5.10 Teaching continued

Teaching continued beyond this group. It continued with more *Gateway* students, separate statistical literacy for arts students and for law students[18], introductory statistics subjects for engineering students, for mathematics and information technology students, and for science students. Teaching continues. Curriculum continues to be developed and activities to illuminate, engage, inspire are still to be found…. The next chapter examines the process through which changes came to be made. It examines the process of reflection. It concludes with a backward and forward looking statement as to those attributes of curriculum and classroom teaching which have been found useful in improving classroom learning.

[18] Porter (1998) examines law students ability to extend their statistical learning to the unfamiliar.

Section 3

On Reflection

I look at my reflection in the mirror
I smile
The Reflection changes.

I smile again
The Reflection changes.

So I frown
The Reflection changes.

I remain neutral
And in doing that too,
The Reflection changes.

I begin to play in the mirror
The mirror in me
Me in the Mirror.

But
Who authors who?

(Norris, 1993, p. 255).

CHAPTER 6

A grounded theory

6.1 The story's end

As the excitement of developing my personal theories of teaching and learning heightened through demonstrable changes in student learning and affect, I began to ask:

- what is it that I have done?
- why did I interpret the literature as I had?
- how were my personal theories regarding teaching and learning modified?
- what was it that led to an improvement in statistical education in my classroom?

My new perspective on theories and methodologies are no longer as they were when I commenced, although often the differences came from subtle shifts in interpretation. In the midst of this heady excitement there was an initial conclusion:

> It was the reflective process led to me to be so profoundly aware of what was happening in my classroom. It was the process that led to changes in teaching practice and curriculum. It was this process that caused me to question established theory. It was through this process that my personal theories of teaching and learning statistics emerged and continue to emerge. (Writings, 1998).

However, throughout the study there was an alarming sense of foolishness as words heard, read and thought understood, took on new meanings, profound meanings, meanings that before the sustained experience of examining the teaching and learning in my classroom were not evident. The plurality of meanings for many terms such as reflective practitioner and teacher-as-researcher and reflection became disturbingly evident.

In this final chapter of the story it has been possible to reflect upon the process and in so doing to reflect upon:

- the theories or work of others with which this study has found common accord and/or discord and associated with this perceived shifts in theoretical perspective by the participant-researcher;
- the question, *What is reflection in the context of relective practice*? and,
- the initial proposition provided by Bain (1990) for improving teaching and learning, that of unpacking the expert in order to better teach the novice;

It has also been possible to reflect upon the product, and to identify:

- those aspects of the teaching experience and improving statistics teaching and learning process that have carried forward to new teaching; and
- areas where further work needs to be undertaken.

6.2 Accord and discord with other theorists: developing theory

The final stages of this work involved revisiting theorists and the writings of teachers-as-researchers in order to identify the similarities and differences with the issues raised in relation to improving education in this study. The purpose of this enterprise was to unite the experiences of the teacher-researcher with those of others and to seek to provide some measure of coherence for the seemingly eclectic theories and practices that were evident in classroom practices in this study. A secondary aim was to illuminate the discord between classroom concerns from this study and those of others. As reflection became interpreted as a process, the areas of discord were viewed as providing templates through which the interpretation of what had happened throughout the study could be re-examined. In the final pass, as it were, through the material the diaries were re-read to draw out the insights that had managed to recede back into the realms of the implicit.

6.2.1 Knowledge is actively constructed by the learner, no matter what pedagogy

Adopting the constructivist proposition that "knowledge is actively constructed by the learner" (von Glaserfeld, 1984) set off a chain of events that shaped the path of this study. In response to this proposition, the teacher-as-researcher set out to develop a set of pedagogical practices, constructivist in orientation, which would require students to generate their own statistical theory. Experiential, problem based learning or activity based learning are often used to exemplify constructivist teaching. Descriptions of the teacher's role in constructivist teaching can be found, for example, the teacher's role is to "provide the setting, pose the challenges, and offer the support that will encourage ... construction"(Davis, Maher, & Noddings, 1990). Rossman in discussing an activity based approach to learning Statistics says:

> These activities require students to collect data, make predictions, read about studies, analyze data, discuss findings, and write explanations...The essential point is that every student is actively engaged with learning the material through reading, thinking, discussing, computing, interpreting, writing, and reflecting. In this manner students construct their own knowledge of statistical ideas as they work through the activities. (Rossman & Chance, 1998, p.132).

This idea of *constructivist teaching* (Gill & Wright, 1994; Jofili & Watts, 1995; Malone et al., 1995) or constructivist learning environments (Garfield, 1995) is common in both the statistical education literature and broader educational field. However as the teacher-researcher examined the reflective homeworks generated by students after the activities, she began to re-interpret the notion of constructivist teaching and meaning of the term reflection and reflective practice. In so doing the first theoretical shift occurred. She concluded that students will construct knowledge, no matter what the environment. There is some supporting literature for this perspective (for example, Klein, 1995). The statement that "knowledge is actively constructed by the learner" is now interpreted as an ontological position, a perspective on how knowledge is formed, rather than a statement of pedagogical practice as implied by the term constructivist teaching. No matter what teaching technique, the pedagogical dictum for the teacher is to determine the nature of students' constructions of knowledge. From this position of knowledge the teacher can determine whether or not the learning needs to be redirected.

Just as Brookfield (1993) recognised that adults do not necessarily know how to engage in 'democratic, critical, authentic, reciprocal, respectful discussion' (p. 25) and experienced embarassment for assuming the ability to discuss was innate, this teacher felt foolish as she discovered that students did not necessarily know how to reflect. In the reflective, homework some students focussed on easy concepts, others on hard, others would be diverted by the theme used to elicit the statistical ideas. Students focussed upon different levels of meaning, for example, in describing data they could distinguish the centre of data as distinct to the spread of data, or that the centre of data could be described by the mean, median and mode, or how to calculate these statistics. Some students described more than one level of meaning. Whilst enmeshed in the implementation phase of the research further theoretical shifts were experienced. The teacher researcher began to realise that:

- her understanding of developmental theories had transmuted and led to the belief that adults would have attained the stage where reflection would simply take place. That is, if they had experience of key statistical ideas through activity, they would know and be able to reflect upon those ideas;

- engaging in activity did not by itself ensure that students would take from the experience that which was to be 'learned'. Evidence of this can be found in the statistical literature, where for example where 'computer simulations alone do not guarantee conceptual change' (delMas, Garfield, & Chance, 1998, p.1083). There needs to be something more. As has been suggested, experience and learning are not synonomous:

  > While experience may be the foundation of learning, it does not necessarily lead to it: there needs to be active engagement with it. (1993, Boud et al p.9).

- The *something more* was in this instance a set of teaching practices guided by the concept of loosening and tightening ideas (Kelly, 1955). Activities were used to loosen thinking. The tightening of ideas achieved through questioning students at the end of the exercise, drawing together ideas and relating them back to the basic statistical notion of statistics is a study of variability, and from the second implementation placing them in the framework of the research process. To have students 'learn' in the sense of being able to apply this comprehensive set of statistical ideas to the solution of some problem and to write about them in the weekly homeworks, involved repetition of the ideas from a variety of perspectives. Each activity drew upon the same set of ideas as the previous activities, but a different focus was formed. The rate at which students came to reflect across the range of ideas, varied.

- mastery of writing about statistics was important if students were to feel confident that they understood. Students who, whilst able to verbally discuss statistical ideas,

> would, at the point of writing, become unsure of themselves and feel that they did not understand the statistics. There are some teachers who use writing as a strategy for teaching statistics. (Grossman, Smith, & Miller, 1993; Radke-Sharpe,1991; Rossman & Chance, 1998; Swanson & McKibben, 1998).

Interpreting what had occurred in this study focussed on the manner in which people reflect. Others interpret differences in reflections as residing not in the reflection itself but within the different experiences of the student. It would be understandable that students might tap into one or other layers of meaning in any experience if:

> [E}xperience … is multifaceted, multilayered and so inextricably connected with other experiences, that it is impossible to locate temporally or spatially. (1993, Boud et al. p.7).

Reflection would yield differences between people if indeed the experience is subject to interpretation:

> The meaning of experience … may not be what at first sight it appears to be. When different learners are involved in the same event, their experience of it will vary and they will construct (and reconstruct) it differently. One person's stimulating explanation will be another's dreary lecture. What learners bring to an event - their expectations, knowledge, attitudes and emotions - will influence their interpretation of it and their own construction of what they experience. (1993, Boud et al, p.11).

It is quite plausible that both what the learner brings to the event and how the student reflects have bearing on the meaning that is derived from an experience. This interpretation is consistent with the teacher-researcher's initial positioning statement, wherein she saw a person's construct system as shaping how events were construed and this approach has not been dismissed. In this study, the teacher-researcher saw reflection as modifiable. (Indeed psychologists do endeavour to modify construct systems). Teacher talk came to encompass both how students needed to reflect and how different layers of meaning could be identified in the material, an idea also reinforced by poetry classes. However, there remains in this summary an emphasis on the notion of there being general patterns of reflection, or perhaps it should be described as modes of responding that divert appropriate reflection. The emphasis is on the output rather than the initial perception of events. The interpretation was that at the point of reflection students were dismissive of various ideas they had encountered and understood. These patterns are problematic for statistics students. Statistics students need to be able to see general patterns, whereas some of these students could only focus on the detail of their own group's data. These characteristic ways of reflecting appeared modifiable. An associated area of study in statistical education would be the study of misconceptions and intuitive thought as applied to particular forms of statistical and probabilistic reasoning.

Students' commentary as to the value of group reflection in the form of *class patter* was in accord with previous finding that:

> … Reflection is not just an individual activity; engaging in the process with another person or with a group can change the meanings we draw from an experience. (Boud, 1993, p.10).

For students the talk of others, students and teacher, was often a stimulus to insight.

The shift in interpretation, recognising that students would construct knowledge no matter what the environment, did not lead to a change in teaching strategy. Students appeared to attain a better understanding of the knowledge than was the case with more traditional approaches (based on the teacher's prior teaching experience). In accord with advocates of experiential learning, it was recognised that:

> Working with our experience is one of the keys to learning… Reflection consists of those processes in which learners engage to recapture, notice and re-evaluate their experience, to work with the experience to turn it into learning. (1993, Boud, pp. 9-10).

The activity based teaching approach together with the associated reflective homeworks provided *a window into the students' learning*. Students responded to the activities; they told stories, made comments, sat in silence, were argumentative; some battled with writing whilst others wrote eloquently. Students became visible to the teacher and their names evoked the image of who they were; others remained invisible. Student responses provided the primary material upon which the teacher-researcher reflected. This, in turn, directed the adjustments to the curriculum and pedagogical approach. This window into student learning is crucial if the learning environment is to be adjusted so as to improve the learning experience.

6.2.2 Reflective practice as research methodology

> A teacher doing action research has the opportunity to pose questions that are relevant to him or her and to seek information that will help the teacher understand his or her own situation more fully. At times, as a result of the research, a different action may be warranted; but at others, no action might be suggested - rather, the teacher-researcher may

> come to understand some facet of education differently. The teacher's research can lead to an enriched or otherwise changed perception of the classroom.(MacGillivray, 1990, p. 188).

Within this rubric of action research there is a multitude of descriptors which define various aspects of the study. The choice of the reflective-practitioner, teacher-as-researcher, or participant-observer mode of research, was in many senses, a pragmatic choice. It permitted a lifestyle approach to research, one that fits, and this for a teacher becomes a necessity if research is to become a matter of course rather than an extra. "The priority [for teachers] is to find a method which works within the constraints of the school day" (Nixon, 1981, p.76). I fulfilled the role of teacher as a researcher of my own practice, and whilst the study involved the study of other experts and learners, it was largely a study of self, directed by the reflective practitioner.

During the implementation phase of the study, if materials collected from students did not lead to a restructuring of the teaching during the subject and did not impact on the current teaching, then the material was set aside. When the act of teaching did not indicate a need to pursue theoretical issues or areas of academic debate, these too were set aside. Through this process there evolved a methodology that led to improved learning. This reflective process was perceived as directing the changes in the classroom and hence as bringing about the improvement.

However, if students did not necessarily know how to reflect, how could I be sure that my reflections upon class events were adequate? This concern moved me beyond the qualitative versus quantitative research debate that replaces validity with viability as the mark of good research. The search for some semblance of truth pervaded all aspects of this study. Lindsey (1999) discusses 'truth' in the statistical and mathematical sense, commenting:

> Models are rather crude simplifications of reality, with epistemological but no ontological status...Most statisticians admit that their models are in no sense 'true', although some argue that there is a true underlying model. (Lindsey, 1999, pp. 3-4).

In this study, while recognising the truth that comes from axioms, there was a different interpretation placed upon truth or lack of truth, not just acceding to the simplification of models. There was a recognition that different statisticians would draw different meanings or implications for definitions and terms. I felt the sense of the whimsical when recently discovering Ford's (1975) classifications of truth, symbolised $Truth_1$, $Truth_2$, $Truth_3$ and $Truth_4$, representing types of truth; basic beliefs, ethical truth, mathematical or logical truth and empirical truth. The search for truth, and later for constructions of truth dominated the unpacking phase and extended into the implementation phase. Even to provide viable accounts, I was suddenly confronted by the question "which voices did I hear in class?" Any teacher embarking upon an action research project faces the problem of when, how and what to record: before, during or after the lesson; emotions, facts, dilemmas or surprises (Enright, 1981, p. 37). To address these dilemmas I engaged in a systematic enquiry into my own practice. Students were photographed and the names of all students learned at the start of the teaching session so that it was easier to identify who had made some impact on the teacher-researcher and who had not. Video-clips of classroom activity were collected. Observations about student learning were taken back to the students for further comment.

I also began to search the literature for studies in reflective practice to see how these researchers addressed the issues of validity. I had perceived myself to be engaged in teacher-as-researcher research, a pursuit that I was finding extremely valuable. Reflective practice was emerging as the unifying theme throughout the study. Where, I began to wonder, were the studies by teachers on their own practice? Recent editorials and articles in *Teachers and Teaching* suggest that teachers' voices should be heard. Denicolo (1996) discussing the International Study Association of Teacher Thinking (ISATT) publications suggests there is:

> ...an abiding concern" that "teachers themselves are given voice in research and scholarship related to teacher thinking. (p. 135).

A similar sentiment is expressed by the editor, Carlgren (1996). That teacher-research is desirable has been often cited:

> [t]eachers are experts about their own context and only they actively can promote their professional development: they can do their own research, they can participate in networks or engage in reflection on their practice. (Wubbels, 1995, p.152)

What I had taken from the collective literature on reflective teaching was the notion that, *I*, the teacher can reflect upon my practice and improve it. *I* am the teacher-as-researcher. *I* engage in my daily teaching, hypothesising and experimenting as *I* attempt to change the educational environment of my students. It is *I* who can give voice to my concerns as a teacher, legitimise the professional knowledge with which *I* work.

Despite the rhetoric in recent years about the need for reflective practice and despite its usefulness in this study, there appeared the possibility that the method was still not considered an appropriate manner of research. Momentarily, my methodology seemed too pragmatic, too opportunistic and, for a statistician, distinctly lacking in design, control and, dare I jest, multivariate analyses. When I returned to the work of Schön (1991) which inspired the reflective practitioner approach adopted, I discovered that only one of the seven case studies was a study of and by the practitioner (or a self study). Others too have noted this absence of accounts by reflective practitioners in the literature:

> With the exception of autobiographies, and stories written by teachers, teachers' voices are still filtered through researchers' writings and publications. Most teachers' voices are co-opted and cannibalized by others; their voices diffused or silenced. For example, in conducting research on teaching, the investigator often interviews, observes, or surveys teachers. But reseachers have their own set of biases and social lenses, their own political views, and often reach their own conclusions - despite what teachers report or say. (Ornstein, 1995, p. 127).

There appeared to be a void in the literature, an absence of teachers' voices in the pursuit of reflective practice in education. Articles giving accounts of teachers reflective practices, as told through another researcher or observer, abound. (See for example, Lucas, 1996). Similarly there are many theoretical discussions or reviews as to what constitutes reflection. (see for example, Bengtsson, 1995; Quicke, 1996; Silcock, 1994; Tsangaridou & Siedentop, 1995). Many articles are devoted to the aquisition and usage of reflective practices by student teachers (Calderhead, 1991; Colton & Sparks-Langer, 1993; Lauriala & Syrjala, 1995). Where were the teachers who gave voice to their own research and in particular to their own reflective practice? Where were the accounts of one's own empowerment as a teacher engaged in reflective practice? Accounts of reflective practice as a deliberate undertaking of the teacher and as reported by the teacher were much rarer than these other accounts.

The outcomes of action research as described by Nixon (1981) included making sense of learners' behaviour, informing the teacher of critical moments in this learning process so that teachers know how and where to intervene, and extended to a modification or elaboration of theories of teaching and learning. These outcomes have each been experienced in this study and, as such, support the use of reflective practice and teacher as researcher methodologies. However to propose a methodology so imbued with the act of reflection, one must address the question "what is reflection?"

6.2.3 What is reflection?

Schön (1983) identified two aspects of reflection practice: reflection-on-action and reflection-in-action. Reflection-on-action encompasses the notion of teachers "revisiting their teaching, perhaps immediately afterwards, but more likely later that same day, or after a much longer period, with more or less attention to evidence, more or less attention to the requirement that such enquiry be systematic" (Lucas, 1996, p. 24). Reflection-in-action is used to describe the process wherein teachers at the chalk-face, as they work, are aware of "the decisions they are taking, and making changes in the light of feedback as they work" (Lucas, 1996, p. 24).

This teacher-researcher had loosely interpreted reflective practice to be "thinking about" what happened in class. This interpretation is found in the literature. When a lecturer and colleague suggested they reflected upon their teaching, modified their practices the next time the course was taught, but still it did not lead to improvements, I began to ask "What have I done in this study?", "How have I reflected?", "What is reflection?" How could I ward off reflection that served to *reify* my existing approaches? (see for example, Smyth, 1992; Zeichner, 1993). As Adler (1993) asked "how do teachers go beyond *mere introspection*, beyond *one-sided subjective, immediate feelings*?" The teacher-researcher came to realise that what had been construed as strategies for getting students to focus upon their own learning had in fact been providing evidence, support and at times disconfirmation of the teacher-researcher's reflections. Reflective practice had become a process for up-ending what was thought known. It became a strategy for looking at practices and student learning from different vantage points. Changes made improved learning outcomes only if the reflection directed the appropriate path and, as I reflected, only if student voices were heard. When changes do not result in an improvement, the changed practice in turn becomes problematic and reflection may suggest another strategy for improvement. This strategy will need to be tested.

The dominant mode of reflection in this study could be described as:

> ...an academic version that stresses reflection upon subject matter and the representation and translation of subject matter knowledge to promote student understanding (Zeichner & Tabachnick, 1991, p. 3).

There may have at times been overtones of other forms of reflection, that which emphasises "the thoughtful application of teaching strategies suggested by research" (called social efficiency) and a developmentalist version emphasising "patterns of developmental growth" (Tabachnick & Zeichner, 1992, p. 3). Reflection did not comfortably extend to the notion evident in the literature of addressing moral, ethical and political aspects of teaching and schooling as discussed by numerous authors and often referred to as critical reflection or as of social reconstructionist persuasion. (Tsangaridou & Siedentop, 1995; van Manen, 1977; Zeichner, 1993). This created a tension throughout the study as I taught students from a variety of disadvantaged backgrounds and during this period I attended conferences imbued with a 'moral, ethical, political and power' discourse. It is not to say that these issues were not reflected upon, but the issues did not acquire potency. Despite Mathematics being described as the gate-keeper subject to a range of professions, reflection within the classroom focussed on how best to improve learning or how best to create a successful learning experience for students from a variety of backgrounds. Within this context, for this discipline and for these students the objective was to attain the standards necessary for entry and subsequent coping at university. Whilst there may have been a recognition of and reflection upon injustice in previous education, it was in relation to an education attained in most cases at least fifteen years prior to undertaking this subject. Zeichner (1993) discussing this tension for action researchers argues that every classroom has a critical dimension in that:

> Teachers can and do make real and important differences in terms of affecting the life chances of their students (p. 201).

In accord with other authors (Rönnerman, 1997) there has been a perception that diary-writing is an important source for reflection. Often-times ideas captured in the diary immediately after teaching would recede, only to attain their prominence upon re-reading. Indeed it was through attempting to write that the researcher became aware of her oscillation between a search for truth and acceptance of many truths. Writing was also found to be crucial for students learning for it was at the point of having to write that students confronted what they had learned or had failed to learn, how learning was organised or their inability to organise the ideas.

While the approach of being a teacher-researcher rather than a traditional researcher was interpreted as having made a profound impact upon, and insight into, the teaching undertaken by the author of this paper, one final reflective comment needs to be made. The process of reflection, teacher-as-researcher or action-researcher, did not necessarily make for easier teaching. Often, it would have been easier to forget what transpired in the classroom than to place one's work under constant critical scrutiny. It also, did not necessarily follow that the learning environment was made safer for students by removing anxiety or uncertainty. Learning to be independent learners and statisticians is a desirable outcome and in this context it is appropriate that students develop strategies to deal with these situations and emotions. This sentiment that it is necessary to have a certain amount of insecurity in mathematics if one were to be a creative or thoughtful user of the discipline was also discovered in the literature. (Krainer, 1993).

Reflective practice had been ascribed the major role in improving statistical education in my classroom. It was seen to be an appropriate form of reflection in that it directed the changes that lead to students' better understanding and more positive affect. However as the euphoria of the reflective-practice experience subsided, I was left to examine Bain's (1990) original proposition and to examine it in the context of there being multiple outcomes embedded in the phrase*, improving statistical education*.

6.2.4 Revisiting Bain: improving understanding

The study upon which this thesis is based was inspired by Bain's (1990) paper regarding how to improve our teaching. The words that captured my attention were those which related to making explicit our implicit knowledge of our discipline. Accompanying this process of making explicit my own (and other experts' ideas) was a search to identify fundamental statistical concepts and how they could be combined. The shift *in student understanding* as measured through end of session grades and concept mapping exercises is attributed to this unpacking of expertise and to repackaging the ideas. When teaching:

- the fundamental statistical idea was kept central to the teaching and learning activity. In this case, the *fundamental idea* was, *statistics is a study of variation* and the *context* for this study was *in the world around us* as. The world around was interpreted in a personal sense of daily thinking and decision making, although the *issues* we might contemplate could be global. As my teaching has moved in the reflective phase, for example, to teaching engineers, this context has been supplemented, not replaced by the context of addressing the *problems and practice of engineering.*

- the ordering of material attempted to both develop and follow students' paths in terms of logical and statistical thinking. Gaps in logic, or teaching units presented, were identified through examining students' assignments and concept maps. As discovered throughout the first implementation, the examination of the ideas needed to be maintained within a cohesive framework or theme. In this case the theme developed was statistics is a study of variation *throughout the research process* (ethics, questions, design, sampling, measurement, description, analysis and drawing conclusions). At times, this theme has been described as a *problem solving path.*

- Within each unit on the research process (questioning, design, sampling,...), a richer more expansive set of statistical ideas arose. The connection between these ideas and variation was established. For example experiements were designed in order to control unwanted sources of variation.

The first two cycles of implementation dealt with establishing and refining the curricula, and this involved a measure of reflection and analysis of my own and others' understanding of the discipline. An improvement in students' understanding was evident in the second cycle; however students' appreciation of what they had learned had declined. The process of making explicit my own knowledge, leading to better explanation and structuring/sequencing of material, was an important part of improving student understanding (or so I reflect). Therefore, the improvement in students' understanding is attributed to a clarification of the ideas that are central to statistics and to creating a flow of logic so that the building blocks as taught were always situated in a holistic perspective.

The statistics discipline itself, needs to address the issue of including the language of variation and variability in the texts it provides for students. In this study the concept of variability was more difficult for students to grasp than sampling, design or the description of data. If "we give prominence to that which we name; and accord it a special place" (Boud, Cohen, & Walker, 1993b, p.14), how then are students to grasp variability as the central statistical concept if student resources are essentially devoid of the term?

6.2.5 Coherence with experiential learning theory

As much of the teaching was based on creating situations where students were exposed to activities, providing statistical and other experiences, it is perhaps appropriate to examine learning from experience or an experiential learning perspective. That is, it is possible to see how the experiences in this study fit within another's theoretical precepts. The set of five propositions about learning from experience developed by Boud, Cohen and Walker (1993a) and the collection of articles on experiential learning in *Using Experience for Learning* (Boud et al., 1993b) can be used to guide this examination. The five propositions are:

1) Experience is the foundation of, and the stimulus for learning;

2) Learners actively construct their experience;

3) Learning is a holistic process;

4) Learning is socially and culturally constructed; and

5) Learning is influenced by the socio-emotional context in which it occurs.

Experience is the foundation of, and the stimulus for learning.

The first of these propositions is explained as follows:

> Experience cannot be bypassed; it is the central consideration of all learning. Learning builds on and flows from experience: no matter what external prompts to learning their might be - teachers materials, interesting opportunities - learning can only occur if the experience of the learner is engaged, at least at some level. These external influences can act only by transforming the experience of the learner. (Boud , Cohen and Walker, 1993, p.8).

The teaching pedagogy utilised in this study was based heavily on providing students with activities or experiences that evoked statistical thought. Students were readily observed, engaged in the physical aspects of completing the exercises. They could be heard generating appropriate statistical thinking. The teaching pedagogy utilised bears strong similarities to that of experiential learning or learning from experience:

> Experience is not simply an event which happens, it is an event with meaning… It is not just an observation, a passive undergoing of something, but an active engagement with the environment, of which the learner is an important part. (1993, p. 6).

The issue of providing for meaningful learning was central to the task of teaching. One interpretation of meaningful learning raised issues as to the sequencing of concepts and an understanding of the structure of ideas.

> Children whose motivation is not very strong soon give up if the subject matter they are asked to learn is not meaningful to them. Meaningful learning occurs only when a student can slot the new material into place among previously learned ideas. If he cannot do this because he has not acquired the necessary subsidiary concepts he can try to remember information (rote learning) but cannot integrate it into his personal thought structure (Gower, 1981, p. 71).

In this approach, the structure of knowledge is fundamental. The use of concept mapping and the thematic approach, (statistics is the study of variation throughout the research process) was considered crucial. The classroom activities were directed toward developing a web of interconnecting ideas. In this study, there two related concepts were identified, that of relevance and authenticity. In this study establishing the relevance of statistics to students' thinking (or discipline) was considered essential. Activities were used to elicit statistical thinking, and this student generated thought formed the basis of the exposition in class. There was however another distinction, that of authenticity. Authentic activity may be defined as the 'ordinary practices of culture'. In this instance this could be viewed as the culture of statisticians. They suggested that:

> 'many of the activities students undertake are simply not the activities of practitioners…this hybrid activity, furthermore limits students' access to the important structuring and supporting cues that arise from the context …The source of such support is often only tacitly recognised by practitioners, or even by teachers' (Brown, Collins, & Duguid, 1996, p. 26).

When teaching statistics, creating both relevance and authenticity was problematic. Statistics is generally a compulsory unit, not selected by the student. In the case studies presented, authenticity was created (or so attempted) through developing connections between statistical thought and everyday decision making, between statistics and students' disciplines. Rarer, but forever sought, was the establishment of authenticity through students' engagement with the discipline itself.

Learners actively construct their experience

The second of these propositions gained salience much earlier in the thesis. To this the teacher-researcher has added the rider "no matter what pedagogical practice is implemented". The experience of the learner in this study was seen to be influenced by their past experiences, their cultural and personal history, as incorporated into their construals of the world. The manner in which students reflected on experience was not simply interpreted as a consequence of perceiving events differently.

Learning is a holistic process

In the context of experiential learning this is a reference to the continuity of all experience:

> the cognitive (concerned with thinking), the affective (concerned with values and feelings) and the conative or psychomotor (concerned with action and doing). (Boud et al., 1993b)P12

In this study, the focus on the curriculum had ensured attention to the cognitive, to students thoughts and constructions of knowledge; the activity basis encompassed the action and doing and working, with student affect an integral component of all implementations.

Learning is socially and culturally constructed

> Critical reflection is required to examine the influences of our values and culture. We cannot easily transcend our taken-for-granted assumptions and see the conceptual baggage we carry with us. The making problematic of the familiar is an important strategy in moving beyond the mental bonds which constrain us...This can take us inwards on the journey of self-exploration to our past and outwards towards engaging with powerful material and social forces of oppression...The most powerful influence of the social and cultural context is that which occurs through language.(Boud et al., 1993b) P14.

Perhaps one of the limitations in this study, as described by Boud et al (1993b) has been the failure to transcend the mental bonds. There was the recognition of the need to build a mathematical language for students. There was the recognition that some subgroups may need a different pedagogical approach. There was no perception of engaging with the powerful material and social forces of oppression. Perhaps this was because the context of the study was that of providing access for a disadvantaged group to higher education. The focus was on making this educative experience successful.

Learning is influenced by the socio-emotional context in which it occurs

Students' subject and teacher evaluations and students' informal commentaries at the end of the second cycle provided an impetus for further reflection; student learning and student understanding were not the only factors to consider when evaluating the educational process.

> Emotions and feelings are key pointers to both possibilities for, and barriers to, learning. Acknowledging them can enable us to significantly redirect our attentions towards matters we have neglected. Denial of feelings is denial of learning...There are two key sources of influence: past experience and the role of others in the present as supportive or otherwise.(Boud et al., 1993b, p. 15).

In this study, the feelings associated with learning were acknowledged throughout the implementations, but more than acknowledgement was required. As was discovered in the second implementation, understanding was almost too easy as students did not seem to appreciate what it was they had learned. Ratings of the subject declined. There was a need to create more *positive attitudes toward the learning of statistics*. It seemed that an explicit focus rather than implicit one, on how they learned to learn statistics was necessary. Students needed to know that 'learning to learn' was one of their greatest challenges. In learning to learn, students became aware that they were learning to deal with uncertainty, learning to write, learning to feel confident and to feel capable. They needed to overcome their lack of self-esteem and often to overcome their negative experiences of learning mathematics. Only in the third implementation, where coping with the feelings associated with learning was afforded a legitimate and *explicit* role in learning, did students respond positively to the learning of statistics.

There is a partial fit between the experiences of this study and the propositions expounded by experiential learning theory. However this story is about developing a model to improve statistical education based on the data generated during this study. Therefore the words to describe the model have a context; they derive from the experience of this study. In concluding the teacher researcher needs only to explore how the model of improving statistical education began to coalesce and how current statistics teaching now proceeds in another context.

6.3 Coalesence: student lives, curriculum, pedagogy and learning

It was at the ISATT '97 conference during a 'talk' that my curriculum and pedagogical approaches to teaching and learning began to coalesce. The scheduled before-dinner talk, was not a talk. The speaker put on some German music, had the audience sit upright in our chairs and proceeded to direct us in a headmistress's commanding voice through an exercise regime. As she did I asked myself*, what is she doing?* and as I too began the dance, *how do I do this?* and finally *why are we doing this?* I was, after all, ready for a talk. These questions may be used to represent the issues associated with improving statistical education and more particularly statistical understanding These are the questions, what, how and why, that need to be addressed in order to create meaningful learning (Novak & Gowin, 1984). They are issues to do with curriculum.

In the asking I moved from *other*, to *I*, to *we* and this characterises the learning issues associated with the learning of statistics. It is a discipline that belongs to the other, the teacher; this is a subject not taken by choice. The *other* understands what needs to be done and when to do it; what may encompass recognising the structure in problems; the choice of appropriate methods, forms of statistical analysis; or the implications of selecting one design over another. The student must find out *how* to do what is required but it is only at the point of understanding the *why* of statistics, what we do in terms of variation, that they can join the community of statisticians.

Just as the speaker engaged the audience in the dance, through tone and manner, there needs to an engagement of the student in the statistics classroom. The pedagogical techniques used in the *Learning Mathematics and Statistics* subject were many and varied, drawing on many theorists. The principle pedagogic approach involved the use of activity based learning, with a heavy reliance on reflective homeworks. The tone and manner of class were important as a variety of pedagogical techniques was used to make the classroom fun, both noisy and quiet, one involving both listening and speaking, challenging but comforting. There were innumerable techniques, building on students existing ideas and experiences, calling into focus their ways of learning and statistical thinking so that they could better understand the world around them and their primary discipline.

The feelings associated with doing the dance are similar to those of learning statistics. Perhaps one may feel silly, unco-ordinated or lacking, perhaps one may refuse to dance, to leave and for some there is no such compunction, the experience is to be enjoyed.

6.4 Model for improving statistical education

The model for improving statistical education, posed in its broadest sense, deals with a variety of outcomes: understanding, different levels of learning, recall and affect. Improvement requires an ongoing methodology for examining what occurs in the teaching and learning process, in this study the methodology of reflective practice. In this study, reflection became a process for systematically gathering and exploring evidence from students. Reflection extended to include the comparison and contrast of this evidence and the teacher's thoughts on issues with evidence and ideas as expressed in the literature. Through this process of reflection, the teacher needed create a window into students' lives, their understanding and affect.

Teaching takes place in a context, that of students lives. Students come to the classroom with life experience, they have a life both within and outside the current learning episode and, whilst learning, they may be endeavouring to compose a future life. The teacher needs to establish pedagogical practices that engage the students in the learning process. Engagement may serve a variety of functions and be accomplished in different ways for different students. In the context of this study, the following aspects of engagement were critical:

- To engage students in learning it may be necessary to establish that the discipline is authentic in terms of the students composing their lives. In this study, teacher talk, the selection of appropriate activities and videos were used to connect statistics with the use of ideas in everyday life and to establish the relevance of statistics to some future profession; this was to create a sense of authenticity for the discipline. In rarer instances, the authenticity was established through an intrinsic interest in the subject matter of the discipline itself.

- The pedagogical practices implemented in this study were successful in engaging students in the learning process. However, as the activity basis was not altered over the course of the study, other pedagogical approaches have not been attempted. These too may be successful. It was considered opinion that traditional teaching undertaken prior to this study was not successful in engendering meaningful learning for most students. It may be that the traditional teaching was simply less satisfactory in providing a window on what students understood and the knowledge they were creating.

- It is an engagement that positions both the student and the teacher as learners drawing upon ideas that arise through the students' activity and discussion. Knowledge in this sense was not immutably predefined. Students in this sense are also the creators of knowledge. In this study, the teacher learned in terms of seeing statistical knowledge differently, and in identifying where there were gaps in logic, either in the presented material or students' interpretation of it. The teacher also learned about the impact of various pedagogical practices on students and self.

- Ultimately, for a student to understand statistics, engagement must leads to the ideas embedded in the curriculum. In particular the engagement must be such that students experience the what, how and *why* of the discipline. For the statistics discipline, the why derives from statistics being a study of variability.

- The pedagogical approach and the associated assessment requirements required that students reflect upon (or engage with) the activities undertaken in class. As was evident from many of the earlier assignments, student reflection upon the statistical ideas was not necessarily a natural talent. To become successful, many students needed to learn how to reflect across the spectrum of statistical ideas and, in many cases, to learn to write in order to express them.

After the loosening of thought through engagement in activity, there needs to be a tightening of the thinking by embedding the ideas within a structure or conceptual framework. Teacher talk, exposition, writing, concept mapping and questioning are all techniques that may be used in this process of establishing links or relationships between ideas.

To improve understanding of statistical ideas firstly necessitates that students be engaged in the learning process. It also involves an identification of those ideas that are fundamental to the discipline and, in this study, as they are expressed in the curriculum. In this study, the teacher researcher came to adopt one idea. Statistics is a study of variation. This provided the 'why' of all that is done in statistics (or the why of all encountered thus far). The importance of this one concept was demonstrated in relation to the research or problem solving process. Improving understanding in this study has been linked to this process of modifying curriculum on the basis of making explicit our understanding of the discipline. This too has involved a reflective stance.

In this study, it was found that the pedagogical approach must include an *explicit* focus on issues associated with learning, in particular the affect (the feelings) associated with learning. In the context of this study, the explicit focus on learning was interpreted as leading to an improvement in student affect. The issues to contend with in this study included, amongst others:

- both previous and current learning experiences, within the classroom and outside;
- coping with uncertainty;
- being able to communicate ideas both verbally and in writing;
- being able to reflect in an appropriate manner and establishing relevance;
- examining the reactions of people around us (are they supporting our learning?); and
- establishing authenticity of the learning in a context where the student is composing a future life.

In reflecting upon all these components of the educational process, the lives of students, pedagogy, curriculum and learning issues the intent is to provide a window into how students are constructing their knowledge and experiencing the learning process. Reflection also extended to students' experiences of the wider academic context, although this did not generate issues of great concern.

6.5 Moving on to new teaching

The Gateway program ceased to exist two years ago. My teaching in the program spanned five years and nine implementations. During that time, I also taught statistical literacy in separate programs to Arts students and to Law students. The sequel has begun as the process of improving statistical education has moved to a variety of classrooms, statistics for engineers, statistics for the natural sciences and to a general introductory program. The context for doing this is one in which the curricula have been established and follow a pattern similar to the professor-expert's textbook and additional chapters, (exploration of data, probability and random variables, and inference) although the speed at which one traverses the material is different. In each instance, there has been large numbers of students (150-300 students).

Within this context:

1) The opening exercise in my statistics class is one that draws attention to learning issues. Just as I developed a learning framework as a teacher in Gateway, when moving to the mainstream lecturing, I also needed to develop a learning framework. This inclusion of a focus on learning also coincided with an improvement in student ratings. Students dance to the music of Billy Joel and in so doing they generate questions. "What is she doing?' As they dance they demonstrate "How do I do it?" and finally they may question "Why are we doing this?" Some students may confront a learning issue, the need for repetition. They may have danced this dance before.

2) The theme 'statistics is a study of variability through the research or problem-solving process' has again been adopted and an exercise similar to that asked of Gateway students has been used (refer Exhibit 5.7).

3) While the curriculum is predefined, all new topics are introduced from the perspective of how we examine variability. For example, when students are introduced to the topic of binomial random variables, the introduction includes the

idea that binomially distributed random variables have known patterns of variation. The process of making explicit the implicit, has moved to making explicit how each technique or piece of theory used relates to variability.

4) Even within the confines of a large lecture theatre, an activity based approach has been adopted. Students are prompted in various ways so that they generate ideas which are elicited (and in many instances anticipated) by the teacher. The teacher - researcher has adopted a 'lecturing' sequence, review (and scaffold) - activity (loosen thinking, generate thought) - debrief (capture student ideas, extend, refine theory and tighten thinking). Activity and the associated experience of statistical thought are treated as important and epistemologically prior to learning.

5) To allow students the freedom from notetaking to participate, two resources have been developed (and used in addition to a textbook). The first is an abbreviated summary and the second a web-based resource. Neither resource is presented as a simple exposition of theory. A socratic approach has been adopted even in the written materials wherein students are asked to respond to some prompt or activity before they link or hyperlink to the ideas generated by other students. In these debriefs, students' ideas are extended and or refined through the introduction of the relevant theory. Whether or not students actually prefer this approach of activity followed by debrief to that of exposition has not been formally tested. However, an improvement in student satisfaction has been observed as exposition in lectures has been increasingly replaced by an activity based approach. Similarly, whether or not students actually learn or understand better than for the expository approach has also not been formally examined;

6) Even in a discipline specific area such as engineering, the primary activity base is constructed in relation to everyday thinking. (For example, a prompt may ask what factors they would consider important in purchasing a car and how they would know that one type of car were preferable to another?). There is still a need to supplement these with discipline specific activities. Establishing both the relevance and authenticity of statistical ideas is considered to be of primary concern.

7) Assessment has two phases. All within session assessment is process orientated. Students in teams collect data and write revealing the processes they followed to collect and analyse data. There is less concern in marking with the accuracy of interpretation, for example aptly describing a set of data as skewed or normal in shape, than there is with ensuring that the student has in fact realised that shape is an important feature of the data that needs examination. There is an emphasis on students explaining what they are doing, why they are doing it and revealing that they can carry out the various tasks. Feedback is given, for all assignments whether graded fail to high distinction, on how to improve. The final end of session assessment requires both process and reasonable interpretation.

8) The key to improving my teaching is to know how students have constructed their knowledge. Observing and questioning students is an integral component of teaching, as is creating the setting so students are prepared to respond to the questions asked or activities provided. Assessment that is open ended rather than tightly defined is also used to identify what students understand.

6.6 Further Issues

Just as the term reflection came into focus, in the concluding phase of this work other terms have come into focus. In this study, a thick description and video capture have been used to describe the teaching. However, many terms used to describe teaching need to be teased apart to capture all the nuances. To focus on the description of one term may disguise the need to focus on another. For example, describing the teaching attribute of asking questions to identify student thought takes the focus off how to get students to respond (particularly in a classroom where responding is seen as non problematic). Similarly, describing the techniques to gain a response, such as asking for a show of hands, may shift emphasis from how a safe environment for responding is created. This in turn may divert the emphasis from an another orientation, identifying questions students themselves would ask (Boulter, 1997). Every action or non-action in a classroom may be purposeful and as teachers we may be aware of some but not necessarily all that we do, and hence we may or may not be able to fully realise the impact of what we do.

One of these expressions is *activity based teaching*. Activities in this study were used to introduce a topic or new statistical concepts. What is activity based teaching? At its most rudimentary usage it was the use of an activity to provoke statistical thinking or in another sense to provide an experience from which learning could flow. Does this mean that any activity will or that it is just to get the students to do something to interact? In this study there was the sense that the activities were to "surprise students by their knowing". They were to generate appropriate statistical thinking although precisely what ideas were to be generated was not obvious to the students at the time of completing the task. The outcomes were not predictable by the students, and not completely predictable for the teacher.

Others provide a commentary on the use of activities. For example Rossman and Chance (1998, p. 132) have students actively engaged through "reading, thinking, discussing, computing, interpreting, writing and reflecting". In the context of statistics, it has been suggested that:

> ...activities for students should be active not passive, asking questions about something in the students' environment and finding quantitative ways to answer the question (Burrill, 1991, p. 52).

There are many such terms to define, teamwork, real data, real problems, collaborative learning, even the "reading, thinking and writing" terms used by Rossman and Chance. The need to focus on the terms, to further define and, refine them, arises when different contexts illuminate the need for clarification. So for example one could ask: This set of real data was received enthusiastically another was not. How do the data sets differ? In this thesis, a description of what transpired has been provided as has some video footage of the classroom to assist in the reader interpreting the meaning of terms.

6.7 In conclusion

As a researcher, I was ever conscious of the subjective nature of reflective practice (objectivity of any research method is problematic). Just as I came to accept that there is no single correct content of or structure for statistical knowledge, I accept that there may be other interpretations of what transpired in my classroom. The greatest challenge for the teacher-as-researcher is to "to find ways of seeing their classrooms from new and surprising angles" (Nixon, 1981, p. 196). Now, when I teach, I have much to reflect upon. Now there are more questions for me to answer about my own practice and my students' learning. In asking "what is it I have done?" I am seeking to up-end what I have done, to look at my teaching in a new way. When I read the literature now, it is to ask "what else could have emerged? And so the questions continue...

References

A

Adler, S. (1993). Teacher Education: research as reflective practice. *Teacher and Teacher Education, 9(*2), 159-167.

Albert, J. H. (1993). Teaching Bayesian statistics using sampling methods and MINITAB. *The American Statistician, 47*(3), 182-192.

Alexander, S., & Hedberg, J. (1994). Evaluating technology-based learning: Which model? In K. Beattie, C. McNaught, & S. Wills (Eds.), *Multimedia in higher education: Designing for change in teaching and learning* (pp. 233-244). Amsterdam: Elsevier.

Alexander, P. A., Schallert, D. L., & Hare, V. C. (1991). Coming to terms: how researchers in learning and literacy talk about knowledge. *Review of Educational Research, 61*(3), 315-343.

Anderson, G. L. (1994). The cultural politics of qualitative research in education: confirming and contesting the canon. *Educational Theory*, 44 (2), 225-237.

Anderson, J. R. (1985). *Cognitive Psychology*. New York: W H Freeman and Company.

Anderson, J. R. (1987). Skill Aquisition: Compilation of weak-method problem solutions. *Psychological Review, 94,* 192-210.

Ausubel, D. P. (1960). The use of advance organizers in the learning and retention of meaningful verbal material. *Journal of Educational Psychology, 51*, 267-272.

Ausubel, D. P. (1968). *Educational Psychology: A cognitive view*. New York: Holt, Rinehart and Winston.

B

Bain, J. D. (1999). Editorial. *Higher Education Research and Development, 18*(2), 165-172.

Bain, J. D. (1990). *Unpacking the expert to teach the novice.* Proceedings of the 8th Australasian Tertiary Learning Skills and Language Conference, (pp. 119-127), Queensland University of Technology, Brisbane.

Bamberger, J. (1991). The laboratory for making things: developing multiple representations of knowledge. In D. A. Schon (Ed.), *The reflective turn: case studies in and on educational practice,* (pp. 37-62). New York: Teachers College Press.

Barnes, M. (1994). Interaction between theory and practice in Bridging Mathematics. *Proceedings of the Australian Bridging Mathematics Network Conference*, (Vol. 1, pp.1-10), Sydney.

Bartholomew, D. J. (1995). What is statistics? *Journal of the Royal Statistical Society, Series A, 158*, 1-16.

Bean, T. W., Sorter, J., Singer, H., & Frazee, C. (1986). Teaching students how to make predictions about events in history with a graphic organizer plus options guide. *Journal of Reading, 29*, 739-745.

Bell, J. A. (1998). Problems in statistics: learning style, age, and part-time students. *Education, 118*(4), 526-528.

Bell, R. (1983). *The spatial interpretation of repertory grids.* Paper presented at the Paper presented at the Australasian Personal Construct Psychology Conference, University of Wollongong.

Bengtsson, J. (1995). What is reflection? On reflection in the teaching profession and teacher education. *Teachers and Teaching: theory and practice, 1*(1), 23-32.

Ben-Zvi, D. (1997). Learning statistics in a technological environment, *Bulletin of the International Statistical Institute ISI '97 Istanbul 51st Session* (Vol. 1, pp. 411-414). Ankara, Turkey: State Institute of Statistics.

Ben-Zvi, D., & Arcavi, A. (1998). Towards a characterization and understanding of students' learning in an interactive statistics environment. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee, & W. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching Statistics* (pp. 647-653). Voorburg, The Netherlands.

Berry, D. A. (1997). Teaching elementary Bayesian statistics with real applications in science. *The American Statistician, 51*(3), 241-247.

Biggs, J. (1988). Approaches to learning and to essay writing. In R. R. Schmeck (Ed.), *Learning strategies and learning styles* (pp. 185-226). New York: Plenum Press.

Biggs, J. (1999). What the student does: teaching for enhanced learning. *Higher Education research and Development, 18*(1), 57-75.

Biggs, J. B., & Collis, K. F. (1982). *Evaluating the quality of learning: The SOLO taxonomy*. New York: Academic Press.

Bilotti-Aliaga, M. (1998). Introduction to Statistics: enhancement through the use of the graphing calculator. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee, & W. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching of Statistics* (Vol. 2, pp. 929-935). Voorburg, Netherlands: ISI Permanent Office.

Bishop, Y., Fienberg, S., & Holland, P. (1975). *Discrete Multivariate Analysis: Theory and Practice*. Cambridge: The MIT Press.

Blackwell, D. (1969). *Basic Statistics*. New York: McGraw-Hill.

Bleicher, J. (1980). *Contemporary hermeneutics: Hermeneutics as method, philosophy and critique*. London: Routledge and Kegan Paul.

Bloom, B. S., Englehart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of Educational Objectives:The classification of educational goals: Handbook 1. The Cognitive Domain*. New York: Longman Green.

Bolton, N. (1991). Cognitivism: a phenomenological critique. In A. Still & A. Costall (Eds.), *Against Cognitivism*, (pp. 103-123). New York: Harvester Wheatsheaf.

Borovcnik, M. (1990). *Curriculum Developments in German Speaking Countries. Proceedings of the Third International Conference on Teaching Statistics* (Vol. 1, pp 84-87). NZ.Voorburg, The Netherlands: ISI Publications in Statistical Education.

Boud, D., Cohen, R., & Walker, D. (1993a). Introduction: Understanding Learning from experience. In D. Boud, R. Cohen, & D. Walker (Eds.), *Using Experience for Learning* (pp. 1-20). Buckingham: Open University Press.

Boud, D., Cohen, R., & Walker, D. (Eds.). (1993b). *Using experience for learning*. Buckingham: Open University Press.

Boulter, C. (1997). Patterns of collaboration between teachers, researchers and pupils in investigating questions. *Paper presented at the 8th Biennial Conference of the International Study Association on Teacher Thinking, Keil, Germany.*

Bowden, J., & Marton, F. (1998). *The University of learning*. London: Kogan Page.

Brew, A., & Boud, D. (1995). Teaching and research: establishing the vital link with learning. *Higher Education, 29,* 261-273.

Brew, A., & Wright, T. (1990). Changing teaching styles. *Distance Education, 11*(2), 183-212.

Brookfield, S. (1993). Through the lens of learning: how the visceral experience of learning reframes teaching. In D. Boud, R. Cohen, & D. Walker (Eds.), *Using Experience for Learning* (pp. 21-32). Buckingham: Open University Press.

Brown, J. S., Collins, A., & Duguid, P. (1996). Situated cognition and the culture of learning. In H. McLellan (Ed.), *Situated Learning Perspectives* . Englewood Cliffs: Educational Technology Publications, Inc.

Bruce, C., & Gerber, R. (1995). Towards university lecturers' conceptions of student learning. *Higher Education, 29*, 443-458.

Brule, J. F., & Blout, A. (1989). *Knowledge Aquisition*. New York: McGraw-Hill.

Brunner, R. B. (1997). Numbers please! The telephone directory and probability. *The Mathematics Teacher, 90*(9), 6-7.

Burge, J. E. (1999). *Knowledge elicitation tool classification.* [online], http://www.cs.wpi.edu/~jburge/thesis/kematrix.html, (7/10/99).

Burghes. D. N. (1993). Teaching Spearman's Rank Order Correlation Coefficient. *Teaching Statistics, 15*(3), 68-69.

Burrill, G. (1991). Quantitative Literacy - Implementation through teacher inservice. In D. Vere-Jones, S. Carlyle, & B. P. Dawkins (Eds.), *Proceedings of the Third International Conference on teaching Statistics* (Vol. 1, pp. 50-55). Voorburg, The Netherlands: International Statistical Institute.

C

Calderhead, J. (1991). Encourage reflective practice in education: an analysis of issues and programs. *Teaching and Teacher Education, 42*(2), 43-51.

Carlgren, I. (1996). Editorial. *Teachers and teaching: theory and practice, 2*(1), 5-9.

Carlson, W. L. (1999). A case method for teaching statistics. *The Journal of Economic Education, 30*(1), 52-55.

Carter, L., & Mougeot, M. (1998). Use of excel in a first course in statistics for mathematics students. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee, & W. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching of Statistics* (Vol. 2, pp. 841-847). Voorburg, Netherlands: ISI Permanent Office.

Castro, A. E., & Cobo, F. T. (1998). Correlation and Regression in secondary school text books. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee, & W. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching of Statistics* (Vol. 2, pp. 671-676). Voorburg, Netherlands: ISI Permanent Office.

Cerrito, P. B. (1999). Teaching statistical literacy. *College Teaching, 47*(1), 9.

Chenail, R. J. (1995). Recursive frame analysis. *The Qualitative Report, 2(*2), http://www.nova.edu.au/sss/QR/QR2-2/index.html.

Chickering, A. W., & Gamson, Z. (1987). Seven principles for good practice in undergraduate education. *AAHE Bulletin, 39*(7), 3-7.

Chizmar, J. F., & Walbert, M. S. (1999). Web-based learning environments guided by principles of good teaching practice. *The Journal of Economic Education, 30*(3), 248-254 .

Chnielewski, T. L., Dansereau, D. F., & Moreland, J. L. (1998). Using common region in node-link displays: the role of field dependence/independence. *Journal of Experimental Education, 66*(3), 197-207.

Ciaccio, A. D. (1998). Hypermedia and WWW for the teaching of statistics. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee, & W. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching of Statistics* (Vol. 2, pp. 953- 959). Voorburg, Netherlands: ISI Permanent Office.

Cicchitelli, G. (1998). The impact of computer programs on the learning of descriptive statistics: the case of DSTATS. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee, & W. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching of Statistics* (Vol. 2, pp. 849-854). Voorburg, Netherlands: ISI Permanent Office.

Clandinin, D. J., & Connelly, F. M. (1991). Narrative and Story in Practice and Research. In D. A. Schon (Ed.), *The Reflective turn: case studies in and on educational practice*, (pp. 258-282). New York: Teachers College Press.

Clayden, E., Desforges, C., Mills, C., & Rawson, W. (1994). Authentic activity and learning. *British Journal of Educational Studies, 42*(2), 163-173.

Cobb, G. W., & Miao, W. (1998). Bears in Space: Activities to introduce basic ideas of design. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee, & W. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching of Statistics* (Vol. 1, pp. 211-217). Voorburg, Netherlands: ISI Permanent Office.

Cobb, G. E., & Moore, D. S. (1997). Mathematics, Statistics and Teaching. *American Mathematics Monthly, Nov,* 801-823.

Cochran-Smith, M., & Lytle, S. (1993). *Inside-outside: teacher research and knowledge*. New York: Teachers College Press.

Colton, A. B., & Sparks-Langer, G. M. (1993). A conceptual framework to guide the development of teacher reflection and decision making. *Journal of Teacher Education, 44*(1), 45-54.

Cookson, P. W. (1987). Closing the rift between scholarship and practice: the need to revitalize educational research. *Educational Policy, 1*, 321-331.

Cox, D. R. (1997). The current position of statistics: a personal view. *International Statistical Review, 65*(3), 261-290.

Creswell, J. W. (1998). *Qualitative Inquiry and Research Design: choosing among five traditions*. Thousand Oaks, California: Sage Publications.

Crowley, M. (1997). Aligning assessment with classroom practices: a promising testing format. *The Mathematics teacher, 90*(9), 706-711.

Cumming, G., & Thomason, N. (1998). Statplay: Multimedia for statistical understanding. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee and W. K. Wong (Ed.), *Proceedings of the Fifth International Conference on Teaching Statistics* (Vol. 3, pp. 947-952). Voorburg, The Netherlands: ISI permanent Office.

Curcio, F. R., & Artzt, A. F. (1996). Assessing students' ability to analyze data: reaching beyond computation. *The Mathematics Teacher, 89*(8), 668-673.

D

Dallal, G. E. (1990). Statistical computing packages: dare we abandon their teaching to others? *The American Statistician, 44*(4), 265-266.

Dansereau, D. F. (1978). The development of a learning strategy curriculum. In H. F. O'Neil (Ed.), *Learning Strategies,* . New York: Academic Press.

Dargahi-Noubary, G. R., & Growney, J. A. S. (1998). Risk - a motivating theme for an introductory statistics course. *The American Statistician, 52*(1), 44-49.

Davidson, J. L. (1982). The group mapping activity for instruction in reading and thinking. *Journal of Reading, 26*, 52-56.

Davies, N., Lees, R., Smith, S., & O'Neill, R. . In L. Pereira-Mendoza, L. S. Kea, T. W. Kee, & W. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching of Statistics* (Vol. 2, pp. 817-823). Voorburg, Netherlands: ISI Permanent Office.

Davis, R. B., Maher, C. A., & Noddings, N. (1990). *Constructivist views on teaching and learning mathematics*. Reston: National Council of Tecahers of Mathematics.

delMas, R. C., Garfield, J., & Chance, B. (1998). Assessing the effects of a computer microworld on statistical reasoning. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee, & W. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching of Statistics* (Vol. 3, pp. 1083-1089). Voorburg, Netherlands: ISI Permanent Office.

Denicolo, P. (1996). Book Review of Teachers Talk about Teaching: coping with change in turbulent times. *Teachers and Teaching Theory and Practice, 2*(1), 135-140

Denson, P. (1992). Preparing posters promotes learning. *The Mathematics Teacher, 85*(9), 723-724.

Dey, I. (1993). *Qualitative data analysis.* London*:* Routledge.

Diamond, C. T. P. (1997). Book Reviews: (1) using semi-dtructured interviews in small-scale research: a teachers guide and (2) Using Observations in small-scale research: a beginner's guide. *Teachers and Teaching: theory and practice, 3*(1), 142-147.

Diaper, D. (Ed.). (1989). *Knowledge elicitation: principles, techniques and applications.* West Sussex, England: Ellis Horwood Limited.

Dietz, E. J. (1993). A cooperative learning activity on methods of selecting a sample. *The American Statistician, 47*(2), 104-108.

Doig, S. (1994). *Pinning the tail on the donkey: the placement of teacher voice in educational research.* Paper presented at the Annual Meeting of the Australian Association of Research in Education, NewCastle.

Dolphin, L. (1992). *Steps in the Scientific Method.* [online], http://www.best.com/~dolphin/, [4/10/99].

Downie, N. M., & Heath, R. W. (1974). *Basic statistical method.* New York: Harper and Row.

Dye, J. (1996). *Socratic method and scientific method.* [online], http://www.soci.niu.edu/~phildept/Dye/method.html, [4/10/99].

Diekhoff, G. M. (1982). Cognitive maps as a way of presenting the dimensions of comparison within the history of psychology. *Teaching of Psychology, 9,* 115-116.

E

Enright, L. (1981). The diary of a classroom. In J. Nixon (Ed.), *A teacher's guide to action research* (pp. 37-51). London: Grant McIntyre.

Entwistle and Ramsden (1983) *Understanding student learning*. London: Croom Helm.

Eraut, M. (1995). Schon Shock: a case for reframing reflection-in-action? *Teachers and teaching: theory and practice, 1*(1), 9-32.

Ericsson, K. A., & Simon, H. A. (1984). *Protocol analysis: verbal reports as data.* Cambridge,MA: The MIT Press.

Eysenck, M. W., & Keane, M. T. (1990). *Cognitive Psychology: a student's handbook.* Hove, East Sussex: Lawrence Erlbaum.

F

Falk, R. (1981). The perception of randomness. In *Proceedings Fifth Conference of the International Group for the Psychology of Mathematics Education*, (pp. 64-66). Warwick, UK.

Fawcett, A. P. (1983). *Constructing Designs:application of interactive elicitation techniques.* Paper presented at the paper presented at the Fourth Australasian Personal Construct Psychology Conference, University of Wollongong, Wollongong.

Federer, W. T. (1991). *Statistics and society: data collection and interpretation.* New York: M Dekker.

Feigenbaum, E. A., & McCorduck, P. (1983). *The Fifth Generation*. London: Michael Joseph.

Feldman, D. S., Hofman, R., Gagnon, J., & Simpson, J. (1987). *Abacus Concepts, StatView II.* Berkley, CA: Abacus Concepts, Inc.

Finch, S., & Cumming, G. (1998). Assessing conceptual change in learning statistics. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee, & W. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching of Statistics* (Vol. 2, pp. 897-903). Voorburg, Netherlands: ISI Permanent Office.

Finzer, W. F., & Erickson, T. E. (1998). Dataspace - a computer learning environment for data analysis and statistics based on dynamic dragging, visualization, simulation, and networked collaboration. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee, & W. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching of Statistics* (Vol. 2, pp. 825-830). Voorburg, Netherlands: ISI Permanent Office.

Fitzgerald, S. (1998). Instructional methods and statistics achievement at the University level: a meta-analysis. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee and W. K. Wong (Ed.), *Proceedings of the Fifth International Conference on Teaching Statistics* (Vol. 3, pp. 1459-1460). Voorburg, The Netherlands: ISI Permanent Office.

Forbes. (1998). Students and Assessment. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee, & W. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching Statistics* (pp. 677-683). Voorburg, Netherlands: ISI Permanent Office.

Ford, J. (1975). *Paradigms and Fairy Tales: an introduction to the science of meanings*. London: Routledge & K Paul.

Freund, J. F., & Simon, G. A. (1992). *Modern Elementary Statistics*. Englewood Cliffs, NJ: Prentice Hall.

Friedman, H. (1972). *Introduction to Statistics.* New York: Random House.

Fisher, K. M. (1990). Semantic Networking: the new kid on the block. *Journal of Research in Science Teaching, 27*(10), 1001-1018.

Flavell, J. H. (1987). Speculation about the nature and development of metacognition. In F. E. Weinhart & R. H. Kluwe (Eds.), *Metacognition, motivation, and understanding*, (pp. 21-29). Hillsdale, NJ: Erlbaum.

Fransella, F., & Banister, D. (1977). *A manual for repertory grid technique*. London: Academic Press.

G

Gagne, R. (1977). *The Conditions of Learning.* New York: Holt, Rinehart and Winston.

Gal, I., & Garfield, J. (Eds.). (1997a). *The assessment challenge in statistics education*. Amsterdam: IOS Press.

Gal, I., & Garfield, J. B. (1998). Aligning assessment with instructional goals. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee, & W. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching of Statistics* (Vol. 2, pp. 773-779). Voorburg, Netherlands: ISI Permanent Office.

Gardner, H. (1985). *The mind's new science: A history of the cognitive revolution.* New York: Basic Books.

Garfield, J. (1995). How students learn Statistics. *International Statistical Review, 63*(1), 25-34.

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

Garner, R., & Alexander, P. A. (1989). Metacognition: Answered and unanswered questions. *Educational Psychologist, 24*, 143-158.

Garven, F., & Reed, M. (1993). Teaching Statistics to Physiotherapy students. *Teaching Statistics, 15*(3), 81-83.

Geertz, C. (1983). *Local knowledge: further essays in interpretive anthropology.* New York: Basic Books.

Gerber, R., & Boulton-Lewis, G. (1998). Teacher's understanding of graphic representations of quantitative information. *Teachers and teaching: theory and practice, 4*(1), 21-46.

Gill, S., & Wright, D. (1994). A hypercard based environment for the constructivist teaching of Newtonian physics. *British Journal of Educational Psychology, 25*(2), 135-146.

Glaser, B., & Strauss, A. L. (1967). *The Discovery of Grounded Theory: strategies for qualitative research.* Chicago: Aldine.

Glass, G. V., & Stanley, J. C. (1970). *Statistical methods in education and psychology.* Englewood Cliffs, NJ: Prentice-Hall.

Godden, G. (1998). Addressing the nurse's need for personalised tuition in a distance offering of introductory service statistics. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee and W. K. Wong (Ed.), *Proceedings of the Fifth International Conference on Teaching of Statistics* (Vol. 3, pp. 1273-1277). Voorburg, The Netherlands: ISI Permanent Office.

Godino, J. D. (1998). Building and experiementing: a model for meaningful instruction in data analysis. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee, & W. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching of Statistics* (Vol. 2, pp. 905-911). Voorburg, Netherlands: ISI Permanent Office.

Gold, P. C. (1984). Cognitive mapping. *Academic Therapy, 19*, 277-284.

Gordon, R. (1994). Keeping students at the center: portfolio assessment at the college level. *The Journal of Experiential Education, 17*(1), 23-27.

Gordon, S. (1994a). Understanding the reluctant learner: some examples from my practice. *Proceedings of the Australian Bridging Mathematics Network Conference* (Vol 1, pp. 11-18), Sydney University.

Goodchild, S. (1988). School pupils' understanding of average. *Teaching Statistics, 10*(3), 77-81.

Gower, D. (1981). The chemistry of the classroom. In J. Nixon (Ed.), *A teachers' guide to action research* (pp. 61-74). London: Grant McIntyre.

Green, D. (1991). Projects and practical work in GCSE statistics. In D. Vere-Jones (Ed.), *Proceedings of the Third International Conference on Teaching Statistics* (Vol. 1, pp. 224-231). Voorburg, The Netherlands: ISI Publications in Statistical Education.

Guba, E. G. (1990). The alternative paradigm dialog. In E. G. Guba (Ed.), *The paradigm dialog*, (pp. 17-30). London: Sage.

Gunawardena, K. L. d. (1998). Introductory Statistics: A cooperative learning approach. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee, & W. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching of Statistics* (Vol. 1, pp. 179-181). Voorburg, Netherlands: ISI Permanent Office.

Gurney, M. (1989). Implementor or innovator? A teacher's challenge to the restrictive paradign of traditional research. In P. Lomax (Ed.), *The Management of Change* (pp. 13-51). Clevedon, UK: Multilingual Matters.

Griffiths, D., Harper, B., Hedberg, J., Fasano, C., & Rayner, J. (1992). *Interactive Multimedia Package for Statistical Literacy*. Application for a 1993 National Teaching Development Grant, University of Wollongong, .

Griffiths, D., & Stirling, D. (1994). *Statistics I: Modelling Variation and Uncertainty*. University of Wollongong: Department of Applied Statistics.

Griffiths, D., Stirling, W. D., & Weldon, K. L. (1998). *Understanding data: principles and practice of Statistics.* Brisbane: Jacaranda Wiley Ltd.

Grossman, F. J., Smith, B., & Miller, C. (1993). Did you say "write" in Mathematics class? *Journal of Developmental Education, 17*(1), 2-8.

H

Hacking, I. (1965). *Logic of Statistical Inference.* Cambridge: Cambridge Univeristy Press.

Hall, R. H., Dansereau, D. F., & Skaggs, L. P. (1992). Knowledge maps and the presentation of related information domains. *Journal of Experimental Education., 61*(1), 5-18.

Hawkins, A. Discussion: Forward to basics! A personal view of developments in statistical education, *International Statistical Review, 65*(3), 280-287.

Hawkins, A., Jolliffe, F., & Glickman, L. (1992). *Teaching Statistical Concepts.* New York: Longman.

Heitzman, W. R. (1980). *Statistics for Business and Economics.* Boston: Allyn and Bacon.

Higgins, J. J. (1999). Nonmathematical statistics: a new direction for the undergraduate discipline. *The American Statistician, 53*(1), 1-7.

Hindle, D. (1993). *GHMD-923 Quantitative Methods*. Wollongong: D Hindle.

Hogg, R. V. (1991). Statistical education: improvements are badly needed. *The American Statistician, 45*(4), 342-343.

Howell, D. C. (1992). *Statistical Methods for Psychology*. USA: Duxbury press.

Howell, D. C. (1995). *Fundamental Statistics for the Behavioural Sciences*. Belmont, California: Duxbury Press.

Huberty, C. J. (1998). Assessment of Student performance in statistics. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee, & W. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching of Statistics* (Vol. 2, pp. 787-792). Voorburg, Netherlands: ISI Permanent Office.

J

Jaeger, R. M. (1990). *Statistics: a spectator sport.* Newbury Park, California: Sage Publications.

Jofili, Z., & Watts, M. (1995). Changing teachers' thinking through critical constructivism and critical action research. *Teachers and Teaching: theory and practice, 1*(2), 213-227.

Jonassen and Tessmer (1996/7). An outcomes-based taxonomy for instructional systems design, evaluation, and research. *Training Research Journal, 2*, 11-46.

Johnston, S. (1993). A case for the 'person' in curriculum deliberation. *Teaching and Teacher Education, 9*(5/6), 473-497.

Johnson, L. W. (1981). Teaching hypothesis testing as a six step process. *Teaching Statistics, 3,* 47-49.

K

Karake, Z. A. (1990). Enhancing the learning process with expert systems. *Computers in Education, 14(*6), 495-503.

Kellaway, F. G. An image from Montaigne. In S. Murray-Smith (Eds.), *The dictionary of Australian quotations*. Richmond: Heinemann.

Kelly, G. A. (1955a). *The Psychology of Personal Constructs: A theory of personality*. (Vol 1). New York: Norton.

Kelly, G. A. (1955b). *The Psychology of Personal Constructs: Clinical Diagnosis and Psychotherapy*. (Vol. 2). New York: Routledge.

Kelly, G. A. (1977). The psychology of the unknown, *New Perspectives in Personal Construct Theory*, (pp. 1-19). London: Academic Press.

Kennedy, K., Olinsky, A., & Schumacher, P. (1990). Using simulation as an integrated teaching tool in the mathematics classroom. *Education, 111*(2), 275-297.

Kjollerstrom, B., & Martensson, M. (1999). Assessment: the key to changing the way we learn. *CAL-laborate, Oct*, 17-20.

Klein, M. (1995). *Towards social justice: re-theorising the "social" in constructivist practice in mathematics.* Paper presented at the Eighteenth Annual Conference of the Mathematics Education Research Group of Australasia (MERGA), Darwin.

Koestler, A. (1978) *Janus*. London: Hutchinson.

Konold, C., Pollatsek, A., Well, A., Hendrickson-Amherst, J., & Lipson, A. (1990). The origin of inconsistencies in probabilistic reasoning of novices. In D. Vere-Jones (Eds.), *Proceedings of the Third International Conference on Teaching Statistics* (Vol. 1, pp. 357-362). NZ.Voorburg, The Netherlands: ISI Publications in Statistical Education.

Korthagen, F. A. J., & Wubbels, T. (1995). Characteristics of reflective practitioners: towards and operationalization of the concept of reflection. *Teachers and teaching: theory and practice, 1(*1), 51-72.

Krainer, K. (1993). Powerful tasks: a contribution to a high level of activity and reflecting in mathematics instruction. *Educational Studies in Mathematics*, 65-95.

Krishnamurty, G. B., Kasovia-Schmitt, P., & Ostroff, D. J. (1995). *Statistics: an interactive text for the health and life sciences*. Boston: Jones and Bartlett.

Kuzma, J. (1992). *Basic statistics for the health sciences*. Mountain View, California: Mayfield.

L

Lackey, J. (1998). Statistics for dummies: how to make SPC fun and (somewhat) easy. *Quality Progress, 31*(7), 112.

Lapin, L. L. (1990). *Statistics for modern business decisions*. San Diego: Harcourt Brace Jovanovich.

Lauriala, A., & Syrjala, L. (1995). The influences of fresearch into alternative pedagogies on the professional development of prospective teachers. *Teachers and Teaching: theory and practice, 1*(1), 101-118.

Layne, B. H., & Wells, J. B. (1990). Development and Application of CAI Modules for use in Statistics and Research Methods. *Journal of Computers in Mathematics and Science Teaching, 10*(1), 35-38.

Lee, C. (1998a). An assessment of the PACE strategy for an Introductory Statistics course. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee, & W. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching of Statistics* (Vol. 3, pp. 1215-1221). Voorburg, Netherlands: ISI Permanent Office.

Lee, T. (1998b). A study on the multimedia courseware for introductory statistics. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee, & W. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching of Statistics* (pp. 585-592). Voorburg, Netherlands: ISI Permanent Office.

Les, J., Maillardet, R., & Cumming, G. (1998). Online explanations for learners: the 'Play it again Sam' facility. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee and W. K. Wong (Ed.), *Proceedings of the Fifth International Conference on teaching Statistics* (Vol. 3, pp. 1317-1323). Voorburg, The Netherlands: ISI Permanent Office.

Lewis, D., O'Brien, D., & Thampapillai, D. (1990). *Statistics for Business and Economics*. Sydney: Harcourt.

Lincoln, Y. S. (1992). Sympathetic connections between qualitative methods and health research. *Qualitative Health Research, 2*, 375-392.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic Inquiry*. London: Sage Publications.

Lindsey, J. K. Some statistical heresies. *The Statistician, 48*(1), 1-40.

Lindley, D. V. (1990). *Statistical Inference.* In D. Vere-Jones (Eds.), *Proceedings of the Third International Conference on Teaching Statistics* (Vol. 1, pp. 33-41). NZ.Voorburg, The Netherlands: ISI Publications in Statistical Education.

Lipson, K. (1998). Using the TI-83 graphics calculator in a liberal arts statistics course. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee, & W. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching of Statistics* (Vol. 2, pp. 937-943). Voorburg, Netherlands: ISI Permanent Office.

Loughran, J., & Northfield, J. (1996). *Opening the Classroom Door: Teacher, Researcher, Learner*. London: The Falmer Press.

Lucas, P. (1996). Coming to terms with Reflection. *Teachers and Teaching: theory and practice, 2*(1), 23-40.

M

MacGillivray, H. L. (1990). *Teaching large classes of science and engineering students.* Paper presented at the Third International Conference on Teaching Statistics, Dunedin, New Zealand.

Mahmud, Z., & Robertson, C. (1998). Developing and testing a teaching model using experiemental design and interview analysis. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee, & W. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching of Statistics* (Vol. 2, pp. 533-540). Voorburg, Netherlands: ISI Permanent Office.

Malone, J., Neale, J., Tsakalos, J., Mailes, B., Smith, M., green, J., & Prasad, K. (1995). *More problem solving ideas for the Mathematics classroom.* Paper presented at the 15th Biennial Conference of the Australian Association of Mathematics Teachers, Darwin.

Marasinghe, M. G., Meeker, W. Q., Cook, D., & Shin, T. (1996). Using graphics and simulation to teach statistical concepts. *The American Statistician, 50*(4), 342-352.

Margenau, H. (1950). The nature of physical reality, cited in J.Van Brakel, (1976), Some remarks on the prehistory of the concept of statistical probability. *Archive for History of Exact Sciences, 16*, 25-28.

Marton, F., Dall'Alba, G., & Beaty, E. (1993). Conceptions of learning. *International Journal of Educational Research, 19*, 277-300.

Marton, F., & Ramsden, P. (1988). What does it take to improve learning?, *Improving Learning: New Perspectives*. Great Britain: Kogan Page.

Marton, F., & Säljö, R. (1976). Approaches to learning. In F. Marton, D. Hounsell, & N. Entwistle (Eds.), *The experience of learning* . Edinburgh: Scottish Academic Press.

Marshall, C. (1990). Goodness criteria: are they objective or judgement calls? In E. G. Guba (Ed.), *The Paradigm Dialog*, (pp. 188-197). London: Sage.

Matlin, M. W. (1994). *Cognition*. Sydney: Harcourt Brace Publishers.

Mattingly, C. (1991). Narrative reflections on practical actions: two learning experiments in reflective storytelling. In D. A. Schon (Eds.), *The Reflective turn: case studies in and on educational practice*, (pp. 235-257). New York: Teachers College Press.

McBride, A. B. (1996). Creating a critical thinking learning environment: teaching statistics to social science undergraduates. *Political Science and Politics, 29*(3), 517-522.

McElroy, E. E. (1971). *Applied Business Statistics: An Elementary Approach*. San Francisco: Holden-Day.

McFarland, T., & Parker, R. (1990). *Expert systems in education and training*. New Jersey: Englewood Cliffs.

McGraw, K., & Harbison-Briggs, K. (1989). *Knowledge aquisition, principles and guidelines*. Sydney: Addison Wesley.

McKechnie, J. L. E. (1957). *Webster's New Twentieth Century Dictionary of the English Language, Unabridged.* New York: The World Publishing Company.

Mevarech, Z. (1983). A deep structure model of students' statistical misconceptions. *Educational Studies in Mathematics, 14,* 415-429.

Micheal, K., Griffiths, D., Porter, A., & Stirling, D. (1993). *The powers of two: How and why should we teach probability in introductory statistics courses?* Paper presented at the Paper presented at the Statistics ’93 Conference, Wollongong.

Miller, A. (1995). Building grounded theory within educational psychology practice. *Educational and Child Psychology, 12*(2), 5-15.

Miller, J. (1999). *Scientific method.* [on line], http://www.slc.qc.ca/sci/scimeth.htm, [4/9/99].

Mitchell, P. (1985). A teacher's view of educational research. In M. Shipman (Ed.), *Educational research: principles, policies and practices* (pp. 81-96). London: Falmer Press.

Mixon, D. (1991). On not-doing and on trying and failing. In A. Still & A. Costall (Eds.), *Against Cognitivism,* (pp. 27-37). New York: Harvester Wheatsheaf.

Moore, D. S. (1991). *Statistics and Contraversies*. San Francisco: Freeman and Company.

Moore, D. (1993). The place of video in new styles of teaching and learning statistics. *The American Statistician, 47*(3), 172-176.

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

Moore, D. S. (1999). Discussion of the paper by Des Nicholls. *Australian and New Zealand Journal of Statistics, 41*(2), 137-139.

Moore, D. S., & McCabe, G. P. (1993). *Introduction to the Practice of Statistics*. New York: W H Freeman.

Morine-Dershimer, G., S Saunders, S., Artiles, A. J., Mostert, M. P., Tankersley, M., S C Trent, S. C., & Nuttycombe, D. (1992). Choosing among alternatives for tracing conceptual change. *Teaching and Teacher Education, 8*(5/6), 471-483.

Mosteller, F., Fienberg, S. E., & Rourke, R. E. K. (1983). *Beginning statistics with data analysis*. Reading, Mass.: Addison-Wesley.

Mosteller, F., & Tukey, J. W. (1977). *Data Analysis and Regression*. Sydney: Addison-Wesley Publishing Company.

N

Naysmith, J., & Palma, A. (1998). Teachers talking, teachers reflecting: how do teachers reflect on their practice? A case study. *Teachers and Teaching: theory and practice, 4*(1), 65-76.

Ng, A., & Nobar, P. M. (1998). Use of commercial computer software in higher education institutions with special reference to statistical software. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee, & W. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching of Statistics* (Vol. 3, pp. 1333-1340). Voorburg, Netherlands: ISI Permanent Office.

Ngee, H. S. L. (1998). Understanding statistical design and analysis of experiments in the classroom. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee, & W. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching of Statistics* (Vol. 1, pp. 219-224). Voorburg, Netherlands: ISI Permanent Office.

Nie, N. H., Hull, C. H., Jenkins, J. G., Steinbrenner, K., & Bent, D. H. (1975). *SPSS: Statistical Package for the Social Sciences.* (2nd ed.). New York: McGraw-Hill.

Nicholls, D. (1999). Statistics into the 21st century (with Discussion). *Australian and New Zealand Journal of Statistics, 41*(2), 127-139.

Nixon, J. (Ed.). (1981). *A teachers guide to action research: Evaluation, Enquiry and development in the classroom*. London: Grant McIntyre Ltd.

Norris, J. (1993). Adulthood ... lost, childhood ... found? *Educational Action Research, 1*(2), 255.

Novak, J. D., & Gowin, D. B. (1984). *Learning how to learn*. Cambridge: Cambridge University Press.

Novak, J. D., & Tyler, R. W. (1977). *A Theory of Education*. London: Cornell University Press.

O

Olson, S. (1999). Candid Camera. *Teacher Magazine, May/June,* 29-32.

Ornstein, A. C. (1995). The new paradigm in research on teaching. *The Educational Forum, 59*(2), 124-130.

Ott, L., Larson, R. F., & Mendenhall, W. (1983a). *Statistics a tool for the social sciences.* Boston: Duxbury Press.

Ott, L., & Hildebrand, D. K. (1983b). *Statistical thinking for managers*. Boston: Duxbury Press.

P

Pask, G. (1988). Learning Strategies, teaching strategies, and conceptual or learningstyle. In R. R. Schmeck (Ed.), *Learning Strategies and Learning Styles* (pp. 83-99). New York: Plenum.

Paul, R. (1992). Critical Thinking: what, why and how in Critical Thinking: Educational Imperative. In C. Barnes (Ed.), *New Directions for Community Colleges, Number 77* . San Francisco: Jossey-Bass.

Payne, R. W., Lane, P. W., Ainsley, A. E., Bicknell, K. E., Digby, P. G. N., harding, S. A., Leech, P. K., Simpson, H. R., Todd, A. D., Verrier, P. J., White, R. P., Gower, J. C., Wilson, G. T., & Paterson, L. J. (1987). *GENSTAT 5 Reference Manual.* Oxford: Clarendon Press.

Peterson, I. (1991). Pick a sample: learning statistics - minus the frustrations. *Science News, 140*(4), 56-59.

Pfannkuch, M., & Brown, C. M. (1996). Building on and challenging students intuitions about probability: can we improve undergraduate learning? *Journal of Statistical Education [online], 4*(1), http://www2.ncsu.edu/ncsu/pams/stat/info/jse/homepage.html.

Phillips, B., & Jones, P. (1991). Developing statistical concepts for Engineering ctudents using computer packages. In D. Vere-Jones (Ed.), *Proceedings of the Third International Conference on Teaching Statistics* (Vol. 2, pp. 255-260). Voorburg, The Netherlands: ISI Publications in Statistical Education.

Piaget, J. (1950). *The Psychology of Intelligence.* London: Routledge & Kegan Paul.

Pike, D. J. (1994). *Statistics in Practice: why is reality so different from expectation?* Paper presented at the Fourth International Conference on Teaching Statistics, Marrakech, Morrocco.

Pinar, W. (1975). The analysis of educational experience. In W. Pinar (Eds.), Curriculum theorizing: The reconceptualists (pp348-395). CA: McCutchan Publishing.

Pope, M. & Denicolo, P. (1993). The art and science of constructivist research in teacher thinking. *Teaching and Teacher Education, 9*(5/6), 529-544.

Pope, M., Denicolo, P., & de Bardi, B. (1990). Professionalism and teacher education. *International Jounal of Personal Construct Psychology, 3,* 313-326.

Porter, A. (1987). *Models for ordered categorical data:* Thesis component, Diploma in Mathematics, Department of Mathematics, University of Wollongong.

Porter, A. (1995). Attaining the balance between learning the Statistics discipline's content and processes, and learning how to learn. Proceedings and Papers, *Bridging Mathematics Network Conference* (pp. 36-48). Batchelor College.

Porter. (1993a). *Unpacking Statistical expert's knowledge in order to teach better.* Paper presented at the Faculty of Education, Postgraduate Education Colloqium, University of Wollongong.

Porter, A. (1996). Focussing on the self as a learner of Mathematics and Statistics. *Technology in Mathematics Education* (pp. 461-468) *MERGA*.

Porter, A. (August 1996). Reflective teacher, reflective students: a case study in improving teaching and learning in Statistics. Proceedings of the *Research in the Learning of Statistics Conference* (pp 77-81). Victoria University, Wellington,

Porter, A. (June 1998a). Statistical Literacy for Law students: six hours to teach. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee and W. K. Wong (Ed.), *Proceedings of the Fifth International Conference on Teaching Statistics* (pp. 533-540). Voorburg, Netherlands: ISI Permanent Office.

Porter, A. L. (1998b). Integrating generic skills into Teaching. *Overview, 5(1), 8-13.*

Porter, A., Griffiths, D., & Hedberg, J. (1993b). *Repacking the expert's knowledge in teaching statistics.* Paper presented at the Statistics '93 Conference, University of Wollongong, Australia.

Porter, A., Griffiths, D., & Hedberg. J. (July 1994). A statistical and mathematical tertiary preparation course, *Proceedings of the Bridging Mathematics Conference.* (Vol 1, pp. 131-138) University of Sydney, Australia.

Porter, A., Griffiths, D., & Hedberg. J. (July 1994). The role of language and experience in the teaching of statistics, *Proceedings of the Bridging Mathematics Conference*. (Vol 1, pp. 139-152) University of Sydney, Australia.

Porter, A., Griffiths, D., & Hedberg, J. (1994c, Dec 5-6). *(Re)constructing knowledge in statistics through reflection, articulation and language.* Paper presented at the Contemporary Approaches to Research in Mathematics, Science and Environmental Education Symposium, Deakin University, Australia.

Portier, S. J., & van Buren, H. A. (1995). An interactive learning environment (ILE) to study statistics: effects of prior knowledge on the use of embedded support devices. *European Journal of Psychology of education, X*(2), 197-207.

Purkey, W. W., & Novak, J. M. (1984). *Inviting School Success*. Belmont: Wadsworth.

Q

Quicke, J. (1996). The Reflective Practitioner and Teacher Education: an answer to critics. *Teachers and Teaching: theory and practice, 2*(1), 11-22.

R

Radke-Sharpe, N. (1991). Writing as a component of Statistics Education. *The American Statistician, 45*(4), 292-293.

Rauch-Hinden, W. (1988). *A guide to commercial artificial intelligence: fundamentals and real-world applications*. Englewood Cliffs: Prentice Hall.

Ramsden, P. (1995). Improving the quality of higher education: lessons from research on student learning and educational leadership. *Legal Education Review*, 6(1), 1-21.

Ramsden, P., Beswick, D., & Bowden, J. (1987). Learning Processes and Learning Skills. In j. T. E. Richardson, M. W. Eysenck, & D. W. Piper (Eds.), *Student Learning: research in Education and Cognitive Psychology* (pp. 168-174). Milton Keynes, England: Open University Press.

Rewey, K. L., Dansereau, D. F., Skaggs, L. P., Hall, R. H., & Pitre, U. (1989). Effect of scripted cooperation and knowledge maps on the recall of technical material. *Journal of Educational Psychology, 81,* 604-609.

Roberts, L. (1993). *Student ownership of data in a lecture-tutorial situation.* Paper presented at the Statistics '93 Conference, Wollongong.

Roberts, L. (1996). *Using concept maps to measure statistical understanding.* Paper presented at the Statistical Education Workshop, Sydney International Statistical Conference, University of Technology.

Robinson, G. (1997). Industrial Statistics. *Statistical Society of Australia Incorporated: Newsletter, 81*, 9-11.

Rönnerman, K. (Oct 1-5 1997). *Diary-writing,. Possibilities for reflective learning.* Paper presented at the Poster No11, presented at ISATT, Keil, Germany.

Roscoe, J. T. (1975). *Fundamental research statistics for the behavioral sciences.* New York: Holt, Rinehart and Winston.

Rossman, A. J. (1994). *Learning Statistics through self-discovery.* Paper presented at the Fourth International Conference on Teaching Statistics, Marrakech, Morocco.

Rossman, A. J., & Chance, B. L. (1998). Workshop statistics: dissemination and assessment. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee, & W. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching of Statistics* (Vol. 1, pp. 131-137). Voorburg, Netherlands: ISI Permanent Office.

Royer, J. M., Cisero, C. A., & Carlo, M. S. (1993). Techniques and procedures for assessing cognitive skills. *Review of Educational Research, 63*(2), 201-243.

Rumsey, D. J. (1999). Cooperative teaching opportunities for introductory statistics teachers. *Mathematics Teacher, 92*(8), 734-737.

Russell, R. J. (1998). Scratching the surface of probability. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee, & W. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching of Statistics* (Vol. 2, pp. 751-756). Voorburg, Netherlands: ISI Permanent Office.

Ryan, S. (1998). Problem-based learning in introductory statistics. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee and W. K. Wong (Ed.), *Proceedings of the Fifth International Conference on Teaching Statistics* (Vol. 3, pp. 1403-1404). Voorburg, The Netherlands: ISI Permanent Office.

Rychlak, J. F. (1983). *George Kelly and the concept of construction.* Paper presented at the Invited Address at the Australasian Personal Construct Psychology Conference, University of Wollongong, Wollongong.

S

Säljö, R. (1979a). Learning about learning. *Higher Education, 8*, 443-451.

Säljö, R. (1979b). Learning in the learners' perspective: some common sense conceptions. *Reports from the Department of Education, 76*(Gothenburg, University Gothenburg).

SAS Institute. (1988). *SAS/STAT User's Guide, Release 6.03 Edition.* Cary, NC:SAS Institute Inc.

SAS Institute. (1989). *JMP User's Guide Version 2 of JMP*. Cary, NC: SAS Institute Inc.

Schacht, S. P. (1990). Statistics textbooks: pedagogical tools or impediments to learning? *Teaching Sociology, 18(*July), 390-396.

Schau, C., & Mattern, N. (1997). Use of map techniques in teaching applied statistics courses. *The American Statistician, 51*(2), 171-175.

Scheaffer, R. L. (1994). *Activity-based Statistics: engaging students in the learning process.* Paper presented at the Fourth International Conference on Teaching Statistics, Marrakech, Morocco.

Schield, M. (1996,). *The goal of Introductory Statistics: Reasoning about data.* Paper presented at the Statistical Education Workshop, Sydney International Statistical Congress (SISC-96), University of Technology, Sydney.

Schloemer, C. G. (1997). Some practical possibilities for alternative assessment. *The Mathematics Teacher, 90*(1), 46-49.

Schmeck, R. R. (1988). An Introduction to strategies and styles of learning. In R. R. Schmeck (Ed.), *Learning strategies and learning styles* . New York: Plenum Press.

Schön, D. A. (1983). *The Reflective Practitioner: How Professionals Think in Action.* New York: Basic Books.

Schön, D. A. (Ed.). (1991). *The reflective turn.* New York: Teachers College Press.

Schuyten, G., & Dekeyser, H. (1998). An electronic learning environment for applied statistics: quality care and statistics education in higher education. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee, & W. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching of Statistics* (Vol. 2, pp. 711-716). Voorburg, Netherlands: ISI Permanent Office.

Searle, S. R. (1989). Statistical computing packages: some words of caution. *The American Statistician, 43*, 189-190.

Seiter, C. (1994). Resampling Stats 1.0 (Resampling Stats Inc.'s statistics/probability software - Software Review). *Macworld, 11*(6), 74.

Seigel, S., & Castellan, N. J. (1988). *Nonparametric Statistics for the Behavioural Sciences.* Singapore: McGraw-Hill Book Company.

Selvanathan, E. A., & Selvanathan, S. (1998). Teaching statistics to business students: making it a success. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee and W. K. Wong (Ed.), *Proceedings of the Fifth International Conference on Teaching Statistics*. (Vol. 3, pp. 1351-1357). Voorburg, The Netherlands: ISI Permanent Office.

Shaughnessy, J. M. (1977). Misconceptions of probability: An experiment with a small-group activity-based, model building approach to introductory probability at the college level. *Education Studies in Mathematics, 8*, 285-316.

Shen, S. M. (1998). Theme-based courses for Tertiary Education. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee, & W. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching of Statistics* (Vol. 3, pp. 1207-1208). Voorburg, Netherlands: ISI Permanent Office.

Shoesmith, E. (1998). Developing a statistics syllabus for finance and accounting. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee, & W. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching of Statistics* (Vol. 1, pp. 183-188). Voorburg, Netherlands: ISI Permanent Office.

Short, T. H. (1997). Discussion. (Response to articles by Donald A. Berry, Jim Albert and David S. Moore). *The American Statistician, 51*(3), 263-265.

Silcock, P. (1994). The process of reflective teaching. *British Journal of Educational Studies., 42*(3), 273-285.

Simon, J. L. (1994). What some puzzling problems teach about the theory of simulation and the use of resampling. *The American Statistician, 48*(4), 290-294.

Singer, J. D., & Willett, J. B. (1990). Improving teaching of Applied Statistics: putting the data back into data analysis. *The American Statistician, 44*(3), 223-230.

Situmeang, R. (1998). A calculator like the TI-85 can speed up the teaching of statistics significantly. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee, & W. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching of Statistics* (Vol. 2, pp. 945). Voorburg, Netherlands: ISI Permanent Office.

Smith, A. F. M. (1996). Mad cows and ecstasy: chance and choice in an evidence-based society. *Journal of the Royal Statistical Society, Series A, 159*, 367-383.

Smith, A. F. M., & Gelfand, A. E. (1992). Bayesian statistics without tears: a sampling-resampling perspective. *The American Statistician, 46*(2), 84-89.

Smyth, J. (1986). *Reflection-in-Action*. Deakin: Deakin University Press.

Symth (1992). Reflective practice in teacher education. *Australian Journal of teacher Education, 17*(2), 11-16.

Sowey, E. (1998). Statistics teaching and the textbook - an uneasy alliance. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee, & W. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching of Statistics* (Vol. 3, pp. 151-158). Voorburg, Netherlands: ISI Permanent Office.

Speed, T. (1986). *Questions, answers and statistics.* Proceedings of the Second International Congress on Teaching Statistics (Vol 2 pp. 18-28), Victoria, Canada.

Spiro, R. J., & Jehng, J. C. (1990). Cognitive flexibility and hypertext: theory and technology for the nonlinear and multidimensional traversal of complex subject matter. In D. Nix & R. Spiro (Eds.), *Cognition, Education, Multimedia: exploring ideas in high technology*. Hillsdale: Erlbaum.

Stedman, M. E. (1993). Statistical Pedagogy: Employing student-generated data sets in introductory statistics. *Psychological Reports, 72*, 1036-1038.

Stenhouse, L. (1975). *An introduction to curriculum research and development*. London: Heinemann.

Stephenson, W. R. (1990). A study of student reaction to the use of minitab in an introductory statistics course. *The American Statistician, 44*(3), 231-234.

Stevens, C. (1994). *Understanding Personal Construct Psychology as Radical Constructivism.* University of Wollongong, Australia: Unpublished Honours thesis.

Stirling, W. D. (1995). *Statistical Exercises using Models'N'Data*. Brisbane: John Wiley & Sons.

Still, A. (1991). Mechanism and Romanticism: a selective history. In A. Still & A. Costall (Eds.), *Against Cognitivism* (pp. 7-26). New York: Harvester Wheatsheaf.

Still, A., & Costall, A. (Eds.). (1991). *Against Cognitivism*. New York: Harvester Wheatsheaf.

Stillings, N. A., Weisler, S. E., Chase, C. H., Feinstein, M. F., Garfield, J. L., & Rissland, E. L. (1995). *Cognitive Science: an Introduction*. Cambridge: The MIT Press.

Stirling, W. D. (1995). *Statistical Exercises using Models'N'Data*. Brisbane: John Wiley & Sons.

Stringer, R. (1993). *Theseus: a strategic approach to bringing about learning*. Paper presented at the ITS Seminar Series, University of Wollongong, Wollongong.

Swanson, D. A., & McKibben, J. N. (1998). On teaching statistics to non-specialists: a course aimed at increasing both learning and retention. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee, & W. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching of Statistics* (Vol. 1, pp. 159-165). Voorburg, Netherlands: ISI Permanent Office.

T

Tooley, J., & Darby, D. (1998) *Educational research : A critique*. Report presented to OFSTED. England: University of Newcastle.

Torrance, E. P., & Rockstein, Z. L. (1988). Styles of thinking and creativity. In R. R. Schmeck (Ed.), *Learning Strategies and Learning Styles* (pp. 275-289). New York: Plenum Press.

Tsangaridou, N., & Siedentop, D. (1995). Reflective teaching: a literature review. *QUEST, 47*, 212-237.

Tukey, J. W. (1954). Unsolved problems of experimental statistics. *Journal of the American Statistical Association, 49*, 706-731.

Tukey, J. W. (1977). *Exploratory Data Analysis.* Sydney: Addison-Wesley.

Turban, E. (1990). *Decision support and expert systems*. Berkeley, CA: Paperback Software International.

Twaite, J. A., & Monroe, J. A. (1979). *Introductory Statistics.* Glenview, Illinois: Scott, Foresman.

U

Ulmer, M. B. (1998). Regression as a foundation for a quantitative elements course for the liberal arts. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee, & W. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching of Statistics* (Vol. 3, pp. 1263-1267). Voorburg, Netherlands: ISI Permanent Office.

University of Wollongong. (2000). *Attributes of a Wollongong Graduate*. Online http://www.uow.edu.au/student/attributes.html. (17/11/00)

V

Valacke, M. (1998). competency-based statistics courses with flexible learning materials. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee, & W. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching of Statistics* (Vol. 2, pp. 723-729). Voorburg, Netherlands: ISI Permanent Office.

Van Brakel, J. (1976). Some remarks on the prehistory of the concept of statistical probability. *Archive for History of Exact Sciences, 61*, 25-28.

van Manen, M. (1995). On the epistemology of reflective practice. *Teachers and Teaching: theory and practice, 1*(1), 33-50.

van Manen, M. (1997). Linking ways of knowing with ways of being practical, *Curriculum Inquiry, 6*(3), 205-28.

van Pattern, J., Chao, C. I., & Reigeluth, C. M. (1986). A review of strategies for sequencing and synthesising instruction. *Review of Educational Research, 56,* 437-471.

von Glaserfeld, E. (1984). An introduction to Radical Constructivism. In P. Watzlawick (Ed.), *The Invented Reality: How do we know what we believe we know?* London: Norton.

von Glaserfeld, E. (1990 An exposition of constructivism: Why some like it radical. In R.B. Davis, C.A. Maher & N. Noddings (Eds.), Constructivist views on the teaching and learning of mathematics. Journal for Research in Mathematics Education, Monograph Number 4, Reston VA:NCTM.

W

Wang, S., & Wong, W. (1998). Teaching statistical analysis for universities; a principled approach. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee, & W. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching of Statistics* (Vol. 2, pp. 593-598). Voorburg, Netherlands: ISI Permanent Office.

Watson, J. M. (1998). Assessment of statistical understanding in a media context. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee, & W. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching of Statistics* (Vol. 2,). Voorburg, Netherlands: ISI Permanent Office.

Weinstein, C. E. (1988). Assessment and training of student learning strategies. In R. R. Schmeck (Ed.), *Learning strategies and learning styles* (pp. 291-313). New York: Plenum.

Wells, A. D., Pollatsek, A., & Boyce, S. (1990). Implications of the NCTM Standards for Mathematics Assessment. In *Mathematics Assessment and Evaluation: Imperatives for Mathematics Educators*. Albany: State University of New York Press.

West, C. K., Farmer, J. A., & Wolff, P. M. (1991). *Instructional Design: Implications from cognitive science*. Sydney: Allyn and Bacon.

Wheatley, G. (1991). Constructivist perspectives on Science and Mathematical Learning. *Science Education, 75(*1), 9-21.

Whimbey, A., & Lochhead, J. (1986). *Problem Solving and Comprehension.* (4th ed.), Hillsdale, NJ: Laurence Earlbaum.

Wild, C. J. (1994). Embracing the "wider view" of statistics. *The American Statistician, 48*(2), 163-170.

Williams, A. L. (1975). *A comparison of the interstimulus onset intervals necessary for the production of apparent movement with visual, auditory, heterogeneous and simultaneous visual and auditory stimuli.* Unpublished Honours Thesis, Department of Psychology, University of NSW (Wollongong Campus), Wollongong.

Williams, A. M. (1998). Students' understanding of the significance level concept. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee, & W. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching of Statistics* (pp. 743-749). Voorburg, Netherlands: ISI Permanent Office.

Winqvist, K. (1991). Statistics in action- learning and comprehending by doing. In D. Vere-Jones (Ed.), *Proceedings of the Third International Conference on Teaching Statistics* (Vol. 2, pp. 546-551). Voorburg, The Netherlands: ISI Publications in Statistical Education.

Witmer, J. (1998). Using Activities in STAT101. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee, & W. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching of Statistics* (Vol. 1, pp. 139-144). Voorburg, The Netherlands: ISI Permanent Office.

Wood, G. R., & Wasimi, S. A. (1998). Transforming first year university statistics teaching. In L. Pereira-Mendoza, L. S. Kea, T. W. Kee, & W. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching of Statistics* (pp. 167-178). Voorburg, Netherlands: ISI Permanent Office.

Wubels. (1995). Editorial. *Teachers and teaching: theory and practice, 4*(1), 149-152.

Y

Yarbrough, S. E., & Gilbert, R. B. (1999). Development, implementation and preliminary assessment of virtual laboratory. *Journal of Professional Issues in Engineering Education and Practice, 125*(4), 147-151.

Z

Zayac, S. A. (1990). *Improving open-ended problem solving: lessons learned in industry. Proceedings of the Third International Conference on Teaching Statistics* (Vol. 1, pp. 79-83). NZ.Voorburg, The Netherlands: ISI Publications in Statistical Education.

Zeichner, K. M. (1993). Action research: personal renewal and social reconstruction. *Educational Action Research, 1*(2), 199-219.

Zeichner, K. M. (1995). Beyond the divide of teacher research and academic research. *Teachers and teaching: theory and practice, 1*(2), 153-172.

Zeichner, K. M., & Tabachnick, B. R. (1991). Reflections on reflective teaching. In B. R. Tabachnick & K. M. Zeichner (Eds.), *Issues and practices in inquiry-oriented teacher education* (pp. 1-21). London: The Falmer Press.

Zidek, J. V. (1988). Statistication; the quest for a curriculum. In D. Vere-Jones (Ed.), *Proceedings of the Second International Conference on Teaching Statistics* (pp. 1-17). University of Victoria, British Columbia, Canada.

Zuwaylif, F. H. (1984). *Applied business statistics*. Boston: Addison-Wesley.