A Study on Teaching Statistical Concepts at the Introductory Level: The Development and Testing of a Teaching Model and an Investigation into the Methodological Techniques

ZAMALIA MAHMUD, MSc

Thesis presented for the degree of
Doctor of Philosophy
in the Department of Statistics & Modelling Science
UNIVERSITY OF STRATHCLYDE
Glasgow
United Kingdom

June 1997

AUTHOR'S DECLARATION

ABSTRACT

An attempt was made to develop a teaching model that could be used to teach basic statistical concepts at the introductory level. The basic framework of the teaching model comprised three teaching components—Provision of Supplementary Notes, Working in Pairs and Teaching Concepts in Worked Examples. These components were identified from a review of studies on teaching methods and an investigational study of students' learning patterns.

The methodological techniques used included observational studies of teaching in lectures, tutorials and practicals, self-report questionnaires on students' learning and teachers' difficulties in teaching statistics, controlled experiments on teaching models, and direct interviews of teachers and lecturers. The experimental studies on teaching methods were conducted at the University of Strathclyde and the interviews took place in schools and universities.

The aim of the main experiment was to test the efficacy of the teaching factors and the effectiveness of the teaching model. The most suitable experimental design was a 2^3 randomised factorial block design with several factors confounded and partially confounded. Eight treatment combinations were tested in small classroom experiments. The main findings indicated the importance of the teaching combination with all three factors present—concepts presented within a worked example, the provision of a supplementary notes and the organisation of students to work in pairs. The implication is that the combination of all three teaching components is required in helping students to learn and it is therefore recommended that this model be further tested in another educational setting.

As well as providing information on the teaching model, this thesis also provides insights on the use of different methodologies. All these methodologies had their advantages and disadvantages but the controlled trials gave the tightest results while the observational studies did not contribute to the testing of the efficacy of the model. Direct interviews have the scope to provide the most detailed information and the use of QDA together with multivariate analysis techniques can help to find patterns among different response patterns.

LIST OF CONTENTS

		page
TITLE		i
AUTHOR'S DECLARATION		ii
ABSTRACT		iii
LIST OF CONTENTS		iv
LIST OF TABLES		ix
LIST OF FIGURES		xiii
ACKNOWLEDGEMENT		xv
CHAPTER 1:	**BACKGROUND OF THE STUDY**	1
1.1	Introduction	1
1.2	The Need for Teaching Statistics	2
1.3	The Teaching of Statistics in Scotland	3
	1.3.1 Statistics in Schools	3
	1.3.2 Statistics at the University: Introductory Level	5
1.4	Current Issues in Teaching Statistics in Scotland	6
1.5	Objective of the Study	8
1.6	Thesis Layout	8
CHAPTER 2:	**RESEARCH INTO TEACHING METHODS**	
2.1	Introduction	11
2.2	Teaching Methods	12
	2.2.1 Lecture versus Discussion	12
	2.2.2 Cooperative Learning in Small Groups	14
	2.2.3 Summary	14
2.3	Teaching and Learning Statistical Concepts	15
	2.3.1 Descriptive Statistics and Probability	15
	2.3.2 Inferential Statistics	19
2.4	Practicals and Projects	21
	2.4.1 Practicals and Computers	23
2.5	Experiments on Teaching Methods in Statistics: The Methodological Issues	24

2.6 Studies on Teaching Methods in Statistics: The Implication of Results 27
2.7 Conclusion 30

CHAPTER 3: A FEASIBILITY STUDY ON CLASSROOM OBSERVATION 33
3.1 Introduction 33
3.2 Classroom Observational Study 34
3.3 Observational Instruments 34
3.4 The Report 35
3.4.1 Lecture Setting 35
3.4.2 Tutorial Setting 36
3.4.3 Laboratory Setting 41
3.5 Students' Learning Patterns 42
3.5.1 Introduction 42
3.5.2 The Instrument 43
3.6 The Analysis of the Questionnaire 44
3.6.1 Students' Opinion on the Lecture Methods 44
3.6.2 Students' Opinion in Different Courses of Study 47
3.6.3 Summary 48
3.7 Conclusion and Discussion 49

CHAPTER 4: LARGE CLASSROOM EXPERIMENT 52
4.1 Introduction 52
4.2 Design of the Experiment 53
4.2.1 Experimental Procedure 53
4.2.2 Teaching Methods 54
4.2.3 The Instrument 55
4.3 Descriptive Analysis 56
4.3.1 Univariate Analysis: Lecture Groups, Gender, Course of Study, Perception 56
4.3.2 Bivariate Analysis 57
4.3.3 Analysis of Key Concepts 58
4.3.4 Analysis of Correct Answers 60
4.4 Conclusion 65
4.5 Analysis of Lecture-Assessment Questionnaire 2 using Logistic Regression Analysis 66

4.5.1 Introduction 66
4.5.2 The Analysis and Results 67
4.6 Conclusion 74
4.7 Discussion 75

CHAPTER 5: SMALL CLASSROOM EXPERIMENT: DESIGN AND METHODS 78
5.1 Introduction 78
5.2 Subjects 79
5.3 The Pilot Experiment 80
5.3.1 Introduction 80
5.3.2 Teaching Methods 80
5.3.3 Procedures 81
5.4 Conclusion 82
5.5 The Main Experiments 83
5.5.1 Introduction 83
5.5.2 Framework of the Teaching Model 84
5.6 Design of the Experiment 85
5.6.1 Introduction 85
5.6.2 Randomisation and Control 86
5.6.3 Statistical Design 87
5.6.4 Instructional Process 89
5.7 Threats to Internal and External Validity 91
5.8 Conclusion 93

CHAPTER 6: SMALL CLASSROOM EXPERIMENT: THE ANALYSIS 96
6.1 Introduction 96
6.2 Methods of Analysis 96
6.3 PHASE 1: Classroom Experiment on 34 students in 1994/95 session 98
6.3.1 Part 1: Descriptive Analysis 98
6.3.2 Logistic Regression Models 100
6.3.3 Checking the Assumption of the Models using Simulation Plots 102
6.3.4 Part 2: Analysis of Model (I) 105

6.3.5 Part 3 - Section (a): Analysis of Model (II) 106
Section (b): Analysis of Model (III) 109
6.3.6 Part 4: Confronting Model (II) and Model (III) 111
6.4 PHASE 2: Classroom Experiment on 24 students in 1995/96 session 113
6.4.1 Part 1: Descriptive Analysis 113
6.4.2 Checking the Assumption of the Models using Simulation Plots 114
6.4.3 Part 2: Analysis of Model (I) 116
6.4.4 Part 3 - Section (a): Analysis of Model (II) 117
Section (b): Analysis of Model (III) 119
6.5 Discussion and Conclusion 122

CHAPTER 7: THE QUESTIONNAIRE SURVEY 125
7.1 Introduction 125
7.2 The Methods 125
7.2.1 Schools and Local Education Authority 126
7.3 The Questionnaire 127
7.4 The Data 128
7.5 Questionnaire Analysis 129
7.6 Discussion and Conclusion 138

CHAPTER 8: THE INTERVIEW : METHODS AND ANALYSIS 142
8.1 Introduction 142
8.2 Procedure and Structure of the Interview 143
8.3 Qualitative Data Analysis (QDA) 144
8.3.1 Conceptual Framework and Research Questions 147
8.4 Response Themes and Codes 149
8.5 Multidimensional Scaling (MDS) and Cluster Analysis 150
8.6 The Analysis of Conceptual Framework 1: Teaching and Learning Difficulties 152
8.6.1 Introduction 152
8.6.2 Learning Difficulties 153
8.6.3 Teaching Difficulties 155
8.6.4 Areas of Misconception 155

8.7 The Analysis of Conceptual Framework 2:
Teaching Methods; their Purpose and Implications 159
8.7.1 Teaching Methods and their Purpose 159
8.7.2 The Implications of the Methods Used 163
8.8 The Analysis of Conceptual Framework 3:
Students' Understanding and Misunderstanding of the Concepts 164
8.8.1 Introduction 164
8.8.2 Evaluation Methods 165
8.8.3 Signposts for Pupils' and Students' Misunderstanding
of the Concepts 166
8.8.4 Signposts for Pupils' and Students' Understanding
of the Concepts 167
8.9 Highlights of the Main Findings 169
8.10 Conclusion and Discussion 169

CHAPTER 9: CONCLUSION AND DISCUSSION 172
9.1 Introduction 172
9.2 Statistical Knowledge 172
9.3 Review of the Development and Testing
of the Teaching Model 173
9.4 Review of the Methodological Issues 177
9.5 The Limitations 179
9.6 Future Research in Teaching Statistical Concepts 180
9.7 Final Words 184

REFERENCES 186

LIST OF APPENDICES 202

LIST OF TABLES

		page
Table 3.1	Teaching Skills Evaluation Form	35
Table 3.2	Observational Outcome of the Tutorial Session (Case Study 1)	37
Table 3.3	Observational Outcome of the Tutorial Session (Case Study 2)	39
Table 3.4	Number of Students who Responded to the Questions	41
Table 3.5	Students' Opinion on Preferred Teaching Methods	44
Table 3.6	How Students Understand Materials Covered in a Lecture?	45
Table 4.1	The Percentages of Students in each Course of Study	56
Table 4.2	The Percentages of Students' Perception on the Lectures	57
Table 4.3	Students in each Course of Study with regard to Lecture Groups	57
Table 4.4	The Male and Female Students in each Course of Study	58
Table 4.5	Students' Perception of the Lecture with regard to Gender	58
Table 4.6	Percentages of Students' Answering Correctly by Lecture Groups	60
Table 4.7	Percentages of Students' Answering Correctly by Course of Study	63
Table 4.8	Percentages of Students' Answering Correctly with regard to their Perception of the Lecture	64
Table 4.9(a)	Chi-square and p-values for both unconditional and conditional effect of Lecture Group for Questions 4 to 12	69
Table 4.9(b)	Chi-square and p-values for both unconditional and conditional effect of Lecture Group for Total p-value and Total c-value Questions	69
Table 4.10	Final Logistic Regression Model for Questions 4 to 12 and the Odds-Ratio of the Significant Variables	71
Table 4.11	Final Logistic Regression Model for Total p-value and Total c-value Questions and the Odds-Ratio of the Significant Variables	73

Table 4.12	Final Poisson Regression Model for the Total Number of Key Concepts Listed and the Odds-Ratio of the Significant Variables	74
Table 5.1	Components of Teaching Models 1 and 2	80
Table 5.2	The Basic Components of the Teaching Model	84
Table 5.3(a)	Experiment Design of Replicate I	88
Table 5.3(b)	Experiment Design of Replicate II (applicable to Phase 1 only)	89
Table 5.4	Teaching Combination in Week 1 (Sampling Distribution of the Means)	90
Table 6.1	Number of times students went to the classes and performed all tests	99
Table 6.2(a)	Students with Correct Answers on the Pre- and Post-tests	99
Table 6.2(b)	The Mean Scores of the Pre- and Post-tests for the eight concepts	100
Table 6.3	Analysis of Deviances for Model (I)	105
Table 6.4	Parameter Estimates for Week	105
Table 6.5	Analysis of Sequential Deviance for Model (II)	107
Table 6.6	Parameter Estimates, Confidence Interval and Odds-Ratio for Model (II)	107
Table 6.7	The Means and Standard Errors for Post-test at Different Levels of Pre-test	109
Table 6.8	Analysis of Sequential Deviance for Model (III)	110
Table 6.9	Parameter Estimates, Confidence Intervals and Odds-Ratio for Model (III)	110
Table 6.10 (a)	Students with Correct Answers on the Pre- and Post-tests (Phase 2)	114
Table 6.10(b)	The Mean Scores of the Pre- and Post-tests for the four concepts	114
Table 6.11	Analysis of Deviances for Model (I)	116
Table 6.12	Parameter Estimates for Week	117
Table 6.13	Analysis of Sequential Deviance for Model (II)	118
Table 6.14	Parameter Estimates, Confidence Intervals and Odds-Ratio for Model (II)	118
Table 6.15	Analysis of Sequential Deviance for Model (III)	119

Table 6.16	Parameter Estimates, Confidence Intervals and Odds-Ratio for Model (III)	120
Table 6.17	The Means and Standard Errors for Post-test at Different Levels of Pre-test	121
Table 7.1	Length of Experience in Teaching Introductory Statistics	129
Table 7.2	Areas of Study at College or University	129
Table 7.3(a)	The proportions and standard errors of the 10 topics most frequently mentioned by the teachers as being difficult to teach	130
Table 7.3(b)	The proportions and standard errors of the 10 topics most frequently mentioned by the teachers as being difficult for pupils to understand	131
Table 7.4	The correlation between the topics most frequently mentioned by the teachers as being difficult to teach and the topics most frequently mentioned by the teachers as being difficult for students to understand (r) and the 95% confidence interval for r	132
Table 7.5(a)	The proportions and standard errors of the 10 topics most frequently mentioned by the lecturers as being difficult to teach	133
Table 7.5(b)	The proportions and standard errors of the 10 topics most frequently mentioned by the lecturers as being difficult for students to understand	134
Table 7.6	The correlation between the topics most frequently mentioned by the lecturers as being difficult to teach and the topics most frequently mentioned by the lecturers as being difficult for students to understand (r) and the 95% confidence interval for r	135
Table 7.7(a)	The percentages of teachers and lecturers who often and sometimes used these teaching methods	135
Table 7.7(b)	The percentages of teachers and lecturers who thought these teaching methods could help to understand the concepts	136
Table 7.8(a)	The percentages of teachers and lecturers who often and sometimes used these teaching materials in their teaching	137
Table 7.8(b)	The percentages of teachers and lecturers who thought these teaching materials could be used to reinforce statistical concepts effectively	138
Table 8.1	Two-way association table for two codes (ABLE and ATT)	151

Table 8.2	The number of teachers and lecturers who mentioned the codes with regard to: (1) pupils' and students' difficulties in understanding and. (2) teachers' and lecturers' difficulties in teaching the concepts	153
Table 8.3	Content analytic summary of teachers' and lecturers' views on pupils and students areas of misconception	156
Table 8.4	Content analytic summary of the teaching methods and purpose of using the methods	160

LIST OF FIGURES

		page
Figure 1.1	Flow Chart of Thesis Layout	9
Figure 4.1	Bar Chart showing the Percentages of Students' Response to Key Concepts in each Lecture Group	59
Figure 4.2	Logistic Regression Model for the Proportion of Correct Answers	66
Figure 4.3	Normal Plot of Pearson's Residual for the Number of Correct Answers	68
Figure 6.1(a)	Simulation (envelope) plot for Model (I) - Phase 1	103
Figure 6.1(b)	Simulation (envelope) plot for Model (II) - Phase 1	103
Figure 6.1(c)	Simulation (envelope) plot for Model (III) - Phase 1	104
Figure 6.2	Response to three factors interaction of NPC (Phase 1)	108
Figure 6.3 (a)	Simulation (envelope) plot for Model (I) - Phase 2	115
Figure 6.3 (b)	Simulation (envelope) plot for Model (II) - Phase 2	115
Figure 6.3 (c)	Simulation (envelope) plot for Model (III) - Phase 2	116
Figure 6.4	Response to three factors interaction of NPC (Phase 2)	121
Figure 8.1	Components of Data Analysis - Interactive Model	145
Figure 8.2(a)	Conceptual Framework 1 - Teaching and Learning Difficulties	147
Figure 8.2(b)	Conceptual Framework 2 - Teaching Methods	148
Figure 8.2(c)	Conceptual Framework 3 - Pupils' and Students' Understanding and Misunderstanding of the Concepts	149
Figure 8.3	2-dimensional solution from MDS showing the relations of all response codes in Matrix Display 1	158
Figure 8.4	A complete linkage dendrogram showing the clustering of all codes in Matrix Display 1	159
Figure 8.5	2-dimensional solution from MDS showing the relations between the response codes on teaching methods	162
Figure 8.6	A complete linkage dendrogram showing the clustering of codes on the methods used in teaching statistics	163
Figure 8.7	2-dimensional solution from MDS showing the relations between the response codes on the methods used to evaluate pupils' and students' understanding/misunderstanding of concepts	165

Figure 8.8 2-dimensional solution from MDS showing the relations between the response codes on the signposts for pupils' and students' misunderstanding of concepts 167

Figure 8.9 2-dimensional solution from MDS showing the relations between the response codes on the signposts for pupils' and students' understanding of concepts 168

ACKNOWLEDGEMENT

First of all, I wish to thank *Allah s.w.t.* for giving me the mental and physical strength, willpower and determination to complete my studies and in putting this novel work in context. The completion of this thesis also would not be possible without continuous and strong support from school teachers, lecturers in higher education and learners of statistics in Scotland.

I wish to thank my sponsors, that is, the Government of Malaysia; Institut Teknologi MARA, Shah Alam, Selangor for awarding me a partial scholarship, and to Overseas Research Students Awards Scheme (ORS Award), Committee of Vice-Chancellors and Principals of the Universities of the United Kingdom and University of Strathclyde for granting me the award towards the completion of my PhD in Statistical Education.

I wish to extend my sincere gratitude and appreciation to my former supervisor, Dr Chris Robertson who was a Senior Lecturer of Statistics at the Department of Statistics and Modelling Science, University of Strathclyde; for his strong guidance, meaningful advice, tremendous patience and endless support throughout my PhD years; for sharing his work experience and in encouraging me to explore the beauty and wonders of statistics in education. Dr Chris Robertson is currently the Deputy Director in the Division of Epidemiology and Biostatistics at the European Institute of Oncology in Milan, Italy.

I would also like to express my thanks and appreciation to Dr Jane O Mitchell, lecturer of Statistics in the Department of Statistics and Modelling Science, Univeristy of Strathclyde for her willingness to continue supervising me in the absence of Dr Robertson. My sincere thanks also goes to Professor George Gettinby, Head of the Department of Statistics and Modelling Science, University of Stratclyde for his professional and fatherly advice and for telling me to keep smiling. My sincere thanks to Mr William Somerville, lecturer of Mathematics and Statistics in Jordanhill College for his meaningful advice as regards the teaching of statistics in Scottish schools and to Mr John McColl, lecturer of Statistics, Glasgow University, for sharing some of his experience in the Statistics in Scottish Schools project. To all my colleagues and lecturers, thank you all very much for the help and support.

Most importantly, I thank God for bringing my family together, through joy and sorrow, in sickness and in health....I owe my success to my beloved husband, Rosli Abdul Rahim who is also pursuing his PhD in Human Resource Management at the University of Strathclyde whom I am constantly praying for success; to my two bundles of joy, Muhammad Dinie Afiq (age 7) and Muhammad Syahmi Amni (age 3) for being so loving, caring and understanding with 'mama' and 'ayah'.

Special thanks is dedicated to my friend Mr Ali Yusob, Mrs Aklah and her children for being very nice to me and my family. Last, but not least, I owe my success to my

caring father Hj. Mahmud bin Kindang, my loving mother Hjh. Tinam bte Mohd Taib and my siblings for their constant prayers and support throughout my whole life.

Chapter 1

BACKGROUND OF THE STUDY

1.1 Introduction

Nowadays, much interest is seen in the teaching of statistics in many disciplines. The interest also arises from the increase in the applications of statistical methods in science and non-science areas. Among them are Biology, Engineering, Business, Economics and Psychology.

Research in educational statistics which deals with issues on methods of teaching, conceptual understanding and reasoning and ways to improve the teaching and learning are among the important areas of research in teaching statistics. The International Conferences On Teaching Statistics (Grey et al., 1983; Davidson, 1988; Vere-Jones David et al., 1991; ISI and NIAC, 1994) have each included a section exclusively on research in teaching statistics. An education section has also been made an important part in the Royal Statistical Society news which offer updates on current activities in statistical education.

The total effort expended on both the teaching and the learning of statistics is enormous and some of the answers are bound up with the nature of statistics and its place in the modern world. According to Anderson and Loynes (1988, pp. 1-2), "Statistics is the study of uncertainty" and in terms of non-probabilistic experience, "Statistics is the collection of ideas and techniques helpful for analysing data."

Chatfield (1988) also mentioned that statistics is concerned with collecting, analysing and interpreting data in the best possible way. He stressed that, a statistician needs to be able to,

- Formulate a real problem in statistical terms.
- Give advice on efficient data collection.
- Analyse data effectively and extract the maximum amount of information.
- Interpret and report the results.

Kendall (1968), on the other hand, differently expressed, "that statistics in its broadest sense is the matrix of all experimental science - and is consequently a branch of scientific method, that it transcends the application of the scientific method in sundry fields of specialisation. The scientist should know statistics as he knows logic and formal language for communicating his ideas. "

Statistics may be thought of as a toolkit which should be widely and wisely used in many disciplines. For example in business and marketing research, statistical methods such as the design of a sample survey is used vastly to investigate the consumers' perception prior to marketing a certain product. As such a good working knowledge in statistics is required in handling such a competent task.

The main issue in this thesis is an attempt to make the teaching of statistics more manageable and understandable that whatever the methods used for teaching statistics should be taken as a challenge by the learner of statistics. There is also a related issue as the investigation of methods to improve statistics teaching leads to an investigation of the methodologies of experimenting on and observing the teaching of statistics. Thus this thesis will also review experimental educational methodologies and this comparison will inform the debate on experiments in education more generally.

1.2 The Need for Teaching Statistics

Many employers feel that newly qualified students entering the job market are not well equipped to practise statistics effectively. Two published reports noted the concern of employers practising statistics. They are: (1) the American Statistical Association (ASA) Committee on the Training of Statisticians for Industry (1980) and, (2) in the United Kingdom, the Joint Working Party of the Royal Statistical Society, the Institute of Statisticians and various groupings of academic statisticians, into the Supply of and Demand for Statisticians (Royal Statistical Society, 1986).

The Joint Working Party found that employers find it difficult to recruit suitable candidates and the cause of it is likely to be due to the teaching of the subject. In that sense, there is the need to develop good teaching practice. Even though the ASA Committee presented guidelines for programmes which train statisticians for industry, not all programmes have yet succeeded in giving adequate emphasis to this aspect. This is where participation in undergraduate teaching and statistical consulting play an important role in the training of prospective teachers of statistics during which they get training in the skills that will make them effective in the classroom. A teacher of statistics also needs to keep up with the changes in the subject matter, advances in instructional technology and new societal demands where competence in these areas can be achieved through in-service training and continued professional development programmes.

Students' learning and practical abilities need to be emphasised in order to prepare them for industry. By understanding and working on real problems, a student learns the art of asking questions, reading relevant material on the topic and gaining experience in the art of communication, criteria which are all important for statisticians and teachers of statistics.

The other problem is the need and neccesity for the subject at school level. This leads on to how marketable is statistics at school: Is there a need for statistics teaching at school and if so, how can this be accomplished.

The issues of trying to diversify the teaching of statistics, to educate people at all levels with statistics, to encourage people to do more research on the teaching of statistics and in the conceptual understanding of the subject were put across in the recent Royal Statistical Society newsletters (RSS News: Vol. 22(10), Vol. 23(1), 1995; RSS News: Vol. 23(8 and 10), 1996).

1.3 The Teaching of Statistics in Scotland

This study is carried out in Scotland and it is pertinent to review the teaching of statistics in that country. Thus where statistics is taught, to whom it is taught and the amount of statistics being taught are investigated. This is accomplished by looking at the general contents of introductory statistics syllabus and discussing briefly the emphasis given by educators who are teaching the subject. This provides background information and establishes the level the thesis is concerned with in the development of the teaching model.

1.3.1 Statistics in Schools

The pupils in Scotland attend seven years of primary school (P1 to P7) from the age of 5 to 11 years. Following that will be a maximum of six years of secondary school (S1 to S6) from the ages of 12 to 17/18. Pupils are required to sit for the National Examination (i.e., Standard Grade) at the end of S4. They can either leave at the end of S4 at the ages of 15/16 or proceed to S5 where they are required to sit for the next National Examination (i.e., Highers) at the age of 17. Many pupils would leave school at the end of S5 to enter Further or Higher Education (McPherson et al., 1990). Provided that they have achieved well in Highers, they may qualify to enter first year university. Others may decided to continue with the Certificate of Sixth Year Studies (CSYS) at S6 or with further highers. At S6 level pupils are offered several courses including CSYS Mathematics I (General paper) which has some statistical contents and CSYS Mathematics III (Statistics paper) with a wider range of statistical contents. The amount of statistics taught at S6 is discussed later in this section.

According to the Education National Guidelines 5-14 Plan (Scottish Education Department, 1991), mathematics should be taught to both the primary and secondary pupils at all schools in Scotland. For many years, statistics has been part of the mathematics teaching in school, both at the primary and secondary level. However, statistics is barely taught at the primary level. At the secondary level, statistics teaching begins to take its course.

There is a growing concern in the teaching of statistics in schools at younger age. The pressure to teach statistics at schools also came about following recent report (Scottish Education Department, 1993) which concludes that 'pupils' attainment in information handling was unduly limited' and that the Scottish Office and Scottish Examination Board should 'take account of the need to include more statistics in courses.'

Some statistical ideas and descriptions are gathered in a textbook, known as the Maths In Action series (Blackie and Chambers; 1989) which is used in many secondary schools in teaching mathematics from between the ages of 12 to 16 in Scotland. A minimum amount of statistical content contributes a section in each series of the textbook. These include a section on basic descriptive statistics and graphs (Book 1, 2 and 3B), elementary probability (Book 4B) and simple statistical survey (Book 5A) .

Unfortunately, it is unlikely that any significant probability and statistical inference will be met until CSYS Mathematics III (Statistics paper) in S6 when pupils are aged 18, that is if the pupils choose to continue to the sixth year level and choose statistics. Little is known how much is being put across effectively by the teachers involved in teaching the subject. It is at the CSYS level that the impact of teaching statistics at schools is at its greatest.

In the revised arrangements in mathematics papers (Scottish Examination Board, 1991), the CSYS Statistics course begins with a review of simple methods of data analysis such as boxplots and stem and leaf diagrams, proceeds through the basics of sampling theory, scatter diagrams and correlation, and concludes with an introduction to hypothesis testing. Interspersed with these statistical topics are the elements of probability, as far as conditional probability and the Binomial, Poisson and Normal distributions. The plan of the course does allow time for experiencing the practical aspects of statistics. An investigative approach also forms an integral part of the teaching process, with class assignments being undertaken to illustrate individual topics. Illustrative examples should be drawn from real-life situations as far as possible.

In 1991, the Scottish Examination Board (SEB) introduced a short course, entitled Introduction to Statistics which is available from S3 onwards. In the Short Course, the pupils are expected to recognise different types of statistical data, to interpret summary measures of location and dispersion, explore relationships between variables using scatter diagrams and two-way tables and to conduct a statistical investigation.

In general, the amount of statistics taught at school from S1 to S4 levels is minimal to such an extent that it provides very little opportunity for pupils to broaden their understanding of the concepts. There is virtually no statistical content at the S5 level. Even though the Maths In Action series used diagrams and illustrations, however, no attempts are made to emphasise the statistical objectives of the exercises. This is one of the issues addressed in this thesis that the process of developing the teaching model will include specifying the statistical objectives of the exercise. This initial investigation into the teaching of statistics in schools in Scotland provides some grounds to look briefly at the amount of statistics taught in the first year university level.

1.3.2 Statistics at the University: Introductory Level

For many years, statistics has had a much favoured position within the degree level Mathematics, Computer Science and Engineering. It is fully integrated into most Mathematics Honours degrees with compulsory courses in the first two years and a wide range of third-year options. However, the amount of statistics taught in these courses will not be emphasised in this section.

Here, the emphasis is on the teaching of introductory statistics courses. The general statistical contents cover a wide range of statistical procedures; concentrating on their correct use; interpretation and clear presentation, applicable to the basic needs of the majority of the students at this level. These include presentation of data, probability and modelling, probability distributions, sampling distributions, simple survey and design and statistical inference.

Introductory statistics is usually taught to students who are required to take statistics in the initial years of their degree program. Among them are students who have not been exposed to statistics at school, at the CSYS level. The syllabus usually starts with descriptive statistics and presentation of data. This is followed by probability and probability distributions and finally the statistical inference section. This is supported by the model of statistics syllabus as mentioned in Hawkins et al. (1992).

In reality, little is known on the actual amount of statistics taught and the emphasis given on the conceptual understanding of the concepts. Judging from the general contents of the syllabus, there is little evidence to suggest that the basic statistical concepts is not covered extensively. However, the extent to which the concepts are emphasised need to be investigated and there is a need for future research in this area.

1.4 Current Issues in Teaching Statistics in Scotland

The purpose is to look at the current state and to explore other developments involving the teaching of statistics. An effort was first made in the Statistics in Scottish Schools project (1994) which purpose was to provide training and to encourage more people to teach and learn statistics.

Since the revised arrangements in mathematics syllabus in 1992, there has been a low uptake in the CSYS Mathematics III (Statistics paper) course among the schools who participated in the Certificate of Sixth Years Studies (CSYS) in Scotland. There has been a slight decrease in the number of schools participating in the CSYS Statistics paper since the introduction of the revised arrangements in mathematics papers. In 1994, 44 schools participated in the revised arrangement for the CSYS Statistics paper (SEB, 1994) as compared to 52 and 48 in 1990 and 1991, respectively, prior to the revised arrangements (SEB, 1990; SEB, 1991). It is also in 1992 that the traditional Statistics paper 3 was replaced by the revised Mathematics III (Statistics paper) at CSYS level and Mathematics I (General paper) modified at the CSYS level.

In the Scottish schools 5-14 plan, information handling and mathematics skills develop very nicely right up to the second year secondary level. It appears that there is a serious lack of progression in teaching statistics from S3 onwards. There is virtually no statistics at all in S5. With the minimum amount of statistics taught in S1 to S4, it is likely that the pupils are going to have difficulties in learning the conceptual aspects of the subject at the university particularly if the subjects were poorly taught. Teachers, when interviewed, stated that they were concerned with the gap which was allowed to exist in the main stream Standard Grade Higher Mathematics. They felt that there is a loss of continuity and this damages the students' perception of statistics at university entry. Then, in 1993 a working party was formed to look at the possibility of including a statistics course in Mathematics Highers (S5 level) and more provision of statistics in Advanced Higher at the Sixth Years Studies level. Finally, the new arrangement in the Mathematics syllabus, known as the Higher Still (SCCC, 1996) was made official to the public and will be implemented in the near future.

Clearly, a considerable amount of effort was put together by the project team towards bridging the gap of statistics teaching and in increasing the provision of statistics at S4 to S6 levels. The aim is to provide a greater awareness on the importance of statistics teaching and learning at schools and in encouraging more educators to teach statistics at school level. However, it is still not known what the future will hold in terms of the educators' perception on the provision of statistics in Higher Still. Obviously, statistics courses are still being offered as an option to the mandatory mathematics courses. Thus, the recognition of statistics courses in school for the university entrance remain doubtful.

Another recent development at the introductory university level is the Teaching and Learning Technology Programme (TLTP catalogue, 1995) which involved the development of the Computer Based Learning (CBL) for statistics such as the STEPS project and the Computer Assisted Learning project (CAL), e.g., QUERCUS.

The STEPS project aims to develop problem based teaching and learning materials for statistics. The problem-based lab materials being produced are based around specific problems arising from Biology, Business, Geography and Psychology. The STEPS project was to provide individual pieces of software which were to be used as resources for students to deal with specific topics. The intention is to motivate and teach students the use of statistics in their own subject area via statistical computer graphic illustrations. On the other hand, the QUERCUS project involved the development of a computer based tutorial material for the Biostatistics course, with very specific purpose—to teach basic statistical techniques, report writing and the use of statistical package Minitab. The course (software) modules were developed for a specific group of students and was built into the course teaching framework. One of the aims of the TLTP projects in statistics is that the computer based learning materials are not intended to replace the lecture but merely for students to use them as backup to the lecture, applicable to their own subject areas. However, much of the work involved in promoting the project was extended at the university level. Thus, knowledge of the work done in STEPS and QUERCUS is not widely known at schools in Scotland.

1.5 Objective of the Study

The main objective of this thesis is in developing and testing a teaching model. This objective requires the use of several methodologies for educational studies. This leads to the second objective namely, to compare and contrast different methodologies of educational experimentation in statistics and in other related areas.

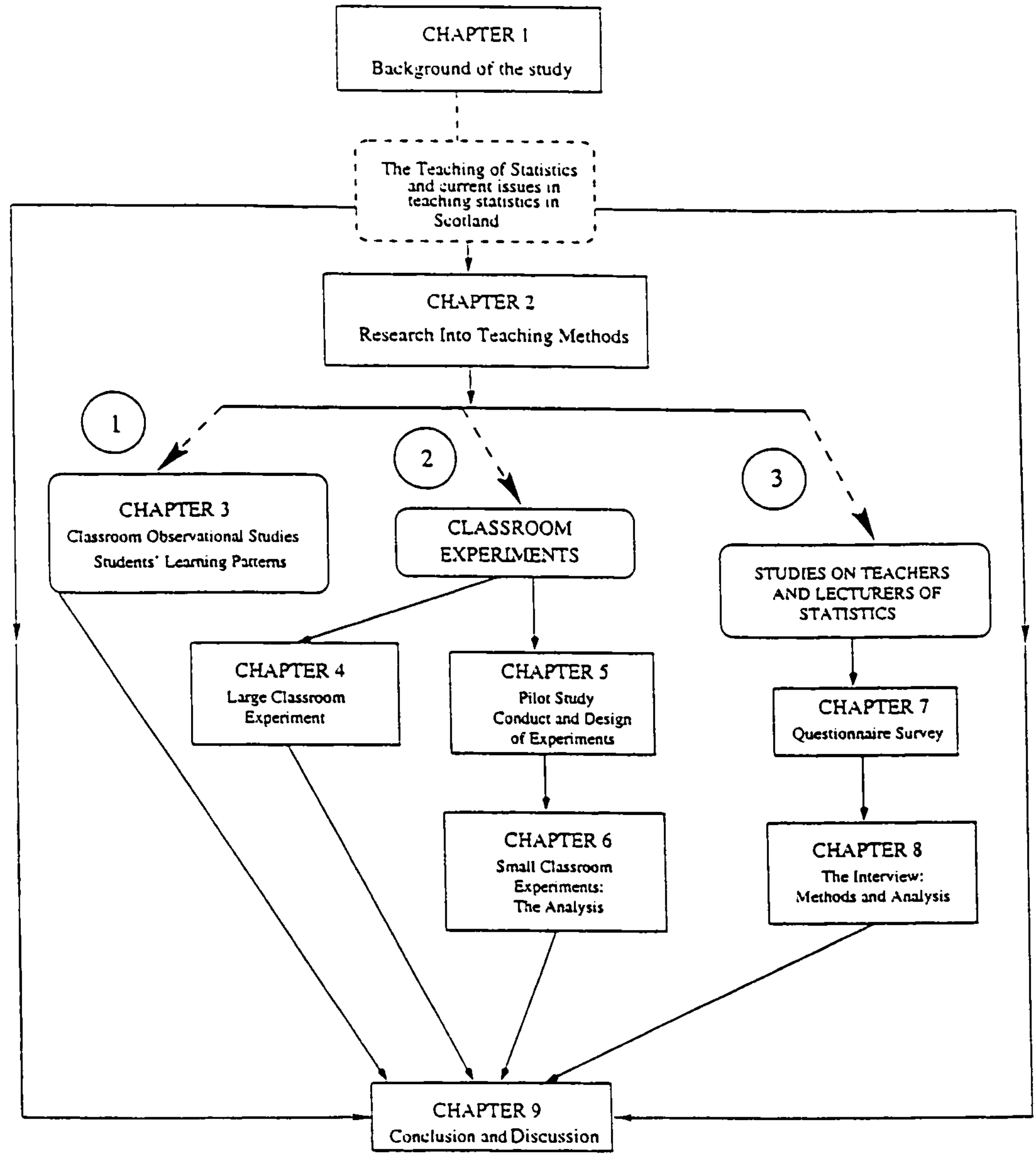
As part of the development of the teaching model, the author used observational techniques, self-report and assessment questionnaires, interviews and comparative experiments in teaching methods. These are the main methodologies which will be examined.

The other aspect of the thesis which is nevertheless important is an investigation into the teaching and learning of statistics. This is not specifically related to the development of the teaching model but this investigation is intended to gain insights on which aspects of teaching and learning of statistics are important to support the choice of the teaching model. This was accomplished by conducting a survey and interview study on school teachers and lecturers in Scotland. In compliance with the second objective of the thesis, several methodological techniques were investigated in the study.

1.6 Thesis Layout

The layout of this thesis begins with a description of a flow chart which briefly describes the contents and the relations between the chapters. The flow chart with the three strands: (1) Classroom Observational Studies, (2) Classroom Experiments and, (3) Survey of School Teachers and Lecturers, respectively, can be seen in Figure 1.1.

Figure 1.1: Flow Chart of Thesis Layout



The whole thesis revolves around teaching statistics at the introductory level in Scotland. Chapter 1 describes the study background. This includes looking at the teaching of statistics at school and university and the current issues in teaching statistics in Scotland.

Chapter 2 is a review of research on teaching methods. This includes methods of teaching, issues on teaching and learning of statistical concepts, studies on teaching methods in statistics, its methodology and design of experiments.

The thesis proceeds with three main strands of educational studies with emphasis given to the second and third strands of the research. The first strand involved initial observational studies (classroom observations of lectures, tutorials and laboratories) and study of students' learning patterns. This is described in Chapter 3. The purpose is to look at the feasibility of conducting experimental studies in educational statistics by observing teaching, investigating lecture styles and students' learning pattern mainly in introductory statistics classes.

The second strand involved some experiments in the classroom. At this point, the experimental studies began to take their course with the first experimental study on teaching methods involving large number of students. These are described in Chapter 4. The study covers several important issues as regards experiments in education, namely that experiments involving large number of students are difficult to control; some measurements and experimental techniques need to be further refined.

Experimental techiques and methods are further explored through a small classroom pilot study with more improvements being carried forward through to the conduct, the design and the analysis of the main experiments. It is at these stages that teaching methods were developed and tested rigorously. All these are described in Chapters 5 and 6.

Then the third and final strand is the study on statistics teachers and lecturers in Scotland. This involved several investigations into the methodologies of conducting research related to the issues governing the teaching of statistics in Scotland. These are described in Chapters 7 and 8.

Chapter 9 conclude the findings of the whole thesis, the contribution to knowledge and implications, the constraints and future research in statistical education.

Chapter 2

RESEARCH INTO TEACHING METHODS

2.1 Introduction

The literature review will focus on issues related to studies on teaching methods. This includes the study design and the methodological techniques in educational studies. The intent is to relate issues in these studies both in statistics and in other areas with the teaching of statistics and experimental investigations into teaching methods in statistics. This includes aspects of teaching and learning statistical concepts and also studies which are concerned with the teaching of statistical concepts. Issues concerning the methodology and the development of other teaching methods shall be discussed in the development of the present study.

There is an enormous literature on teaching methods and it is not the purpose of this chapter to cover all methods. The main aim of this thesis is the development and testing of a statistics teaching model and the literature review retains this focus. An aim is to develop a model with wide applicability among current students and with current resources.

As far as the teaching of statistics is concerned, the principal teaching methods in current use are lecturing, tutorials in small groups or larger groups, individual or group investigations, projects or practicals. Recently computer based tutorial instruction has become available. Undoubtedly its impact will be great but it will not be the main focus of this thesis as the choice has been made to concentrate on teacher (or person) based teaching methods. Also much of the computer based material in the United Kingdom was developed concurrently with this thesis.

The literature review will begin with a very brief discussion of two of the most important methods——lectures and cooperative learning. Following this is a look at the issue on the teaching of basic statistical concepts in Section 2.3 and one important teaching method, the practicals in Section 2.4.

This chapter concludes with a discussion of the methodological implications of previous experiments in teaching statistics. This has a bearing on the experimental methodologies used in this thesis. Finally the implications of the results of these studies as regards the development of the teaching model are debated.

2.2 Teaching Methods

2.2.1 Lecture versus Discussion

One of the dominant modes of teaching continues to be the lecture (Bligh, 1972; McLeish, 1976; Waggoner, 1984; Behr, 1988). Bligh (1972) and Behr (1988) mentioned that a lecture is used for presenting information in an organised way, for identifying or clarifying problems or issues, for presenting the analysis of a controversial issue, for stimulating or inspiring the audience and for leading them on to further study. Woods (1984) mentioned that lectures are especially valuable for introducing and opening up a subject and students can thus be led into subjects which would otherwise prove too daunting for them. Behr (1988) pointed that a lecture can be given its maximum value if the subject matter is well prepared and presented in an attractive manner.

One could argue that styles of lecturing and experience may influence the effectiveness of one's lecture. Behr (1988) stated that there exist lecturing styles which are uninfluenced by experience. At the same time, Brown and Bakhtar (1988) identified five distinctive lecturing styles which are not influenced by experience but associated with subject areas. The five lecturing styles are summarised. The Exemplary style identified lecturers who organised lectures logically and according to a set of objectives. They are usually confident, able presenters and have well-structured lectures. The Amorphous style identified lecturers who are confident and felt to have structured their lectures well but actually do not have selected teaching objectives. The Visual style identified lecturers who structure their lectures well but provide too much detail in a single lecture. The Self-Doubters identified lecturers who have difficulty in selecting and structuring material for their lectures. Finally, the Oral presenters rarely use any means of communication other than talk. Exemplary presenters were found to be more common in the engineering and the pure sciences. Oral presenters were found to be more common in the social sciences and humanities. The Amorphous, Visual and Self-Doubters appear to be spreading across all subject areas. In essence, Exemplary presenters are the best at providing clear, well-structured explanations and to a lesser extent interesting explanations. Well structured explanations tend to yield higher measures of recall and understanding and good explanations have clarity of structure (Gage and Berliner, 1979; Brown and Amstrong, 1984; Brown and Atkins, 1988).

On the other hand, McLeish (1968a) pointed out that knowledge gained by passive listening without some participation by the class is not readily asssimilated and saturation sets in very quickly. It can be argued that a lecture can motivate students if the presentation and organisation of the lecture is carefully planned and implemented to the extent of an Exemplary style. Costin (1972) found that discussion was especially advantageous for promoting retention of factual knowledge and in solving problems. In contrast, Solomon et al. (1964) concluded that the more clearly the teacher expresses themselves, the greater were their students' gains in factual knowledge. Earlier Bloom (1953) discovered that discussion is probably more effective than lectures for teaching cognitive skills, such as interpreting knowledge and solving problems and for helping students retain information beyond the termination of the course. When the criterion of achievement is the ability to solve problems, increasing the emphasis on discussion seems to promote greater learning. This is supported by Dees' (1991) study that the resulting increase in students' problem-solving abilities justified the use of the cooperative method which has the elements of discussion integrated to it.

Woods (1984) explained that in a lecture in which the emphasis is on the transmission of information, it may not be possible for much to be understood during the time. McLeish (1968a) discovered that material presented at the beginning and end of a talk was better remembered than the central sections no matter what information was presented in the three sections. Lloyd (1968) also found that there was an initial increase followed by a decrease in the number of facts noted by students until the last ten minutes of the lecture.

In the above, some of the researchers discovered that passive conventional teaching alone would not be effective in helping students to learn. The method itself either needs to be improved or reformed through activities that would stimulate ideas and enhance students' understanding of the subject matter. They may also be complemented by the use of active learning methods such as the use of small group tasks and discussion groups. The time alloted to presenting materials in a lecture is also important. Hence, materials should be carefully selected and presented appropriately to suit the time. It seems that the elements of discussion integrated into the lecture are an important combination useful for promoting greater learning.

These support the practice of including both lecture and discussion methods in the development of the teaching model.

2.2.2 Cooperative Learning in Small Groups

Cooperative learning was found to be a useful method for stimulating thought and interest in a subject. It consists of a group putting together its knowledge, ideas and opinions in a cooperative manner. Dietz (1993) found that it stimulates students to think creatively about certain problems and help with familiarisation of study materials (Dees, 1991). Webb (1982) stated that students learn better and retain more if they engage in cooperative learning activities. Cooperative learning is useful when the teacher wants to assess whether the group shares a common understanding on the subject and when the group needs to plan to meet its own learning needs.

As pointed by Garfield (1993), the disadvantage of cooperative learning is that it can fail to get going or become irrelevant or heated if the cooperative rules are not clearly stated. This could be rectified if the rules are clearly stated and the responsibility are shared among the group members. Dietz (1993) supported by saying that students who are given the chance to be actively involved in their own learning tend to learn better and gain confidence in their own abilities and intuition . The key issue is working cooperatively towards achieving certain goals. The positive aspect of cooperative learning is that it encourages participation and stimulates thinking which can lead to more understanding of the subject matter. The fact that cooperative learning is found to be successful in some studies has influenced the decision to choose the concept of working together in the development of the teaching model.

2.2.3 Summary

It is recognised that the review in the preceding sections does not comprise a thorough review of lecturing and small group methods. Based on the researcher's personal experience of teaching and of observing others teach, these two methods are among the principal teaching methods which are convenient and fairly easy to implement in the classroom. Thus it was anticipated that they would form the basis of the teaching model. Consequently, the review was intended to deliberately investigate the strengths and the weaknesses of the methods to further support the choices.

The lecture method will continue to be popular in spite of its drawbacks. In their study, Gibbs and Harland (1987) found that the most dominant mode of teaching in higher education is the lecture even though some lecturers resent the fact that it is effective. The positive aspect of a lecture is that it is useful for introducing and providing knowledge of subject to a large group. There is no doubt that the elements of passive learning results in the amount of knowledge received not being assimilated and that there is little indication of how much learning is taking place (McLeish, 1968a). Thus, teaching methods that involve integration of active learning should be greatly considered in the development of any teaching model. Learning is further enhanced by introducing methods that encourage students to work together, help to stimulate their thinking and give the chance to expose their own ability in problem solving. As a result, the type of teaching model chosen for the study is one that is based on lecturing, has elements of cooperative learning and should have units of new material packaged up into small time periods.

2.3 Teaching and Learning Statistical Concepts

As part of the development of a teaching model, the process will include using the teaching model to teach statistical concepts. Issues related to research in teaching and learning of statistical concepts at the introductory level which include the methods used in teaching statistical concepts and the development of teaching methods in teaching statistics will be reviewed. The review will focus on the teaching of descriptive statistics and statistical inference concepts rather than on probability. The reason being that aspects of probability and probabilistic reasoning have already been explored extensively (Fischbein, 1975; Pollatsek et al., 1987; Falk et al., 1980).

2.3.1 Descriptive Statistics and Probability

There are many reasons that most students fail to enjoy introductory statistics and find it hard to learn. Three of the most compelling reasons as stated by Watts (1991) are: (1) Introductory statistics involves concepts which are abstract, (2) Statistical notation and terminology are often ambiguous and confusing, and (3) Statistics problems often have more than one solution, which requires interpretation.

Watts (1991) noted that students must not only grapple with truly abstract concepts but they must immediately relate and apply these concepts to reality. Kalton (1973) and Jolliffe (1976) reported the comments made by instructors about students not attaining an adequate understanding of basic statistical concepts and not being able to solve applied statistical problems. One could suggest that there are several ways of overcoming these problems. Kalton (1973) suggested that students should be equipped with a basic numeracy and mathematical knowledge as this could help to increase their motivation in learning statistics. One could argue that statistical concepts can be taught without relying heavily on students' mathematical knowledge and ability. Thus, Brewer (1994) avoided using number-crunching problems where he quoted that, "The concept of mean, for example, can be grasped and explained with three single-digit numbers as effectively as with a thousand four-digit numbers." It appears that the current trend in developing introductory statistics courses has also moved away from being mathematically based to a more conceptual. integrated approach (Roiter and Petocz, 1996; Yilmaz, 1996).

Students' mathematical ability is usually taken as a guide to their understanding of probability. Research on this has developed a separate area which focuses on the difficulties that students have in understanding probability concepts. A review of the literature confirmed that students appear to have difficulties developing correct intuition about fundamental ideas of probability for at least three reasons:

(1) Students tend to have difficulty with rational number concepts and proportional reasoning, which are used in calculating, reporting and interpreting probabilites (Behr et al., 1983). Students were generally weak in rational number concepts and had difficulties with basic concepts involving fractions, decimals and percents (Carpenter et al., 1983),

(2) Probability ideas often appear to conflict with students' experiences and how they view the world (Kapadia, 1985) and,

(3) Many students have already developed a distaste for probability through having been exposed to its study in a highly abstract and formal way (Garfield and Ahlgren, 1988).

Garfield and Ahlgren (1988) provide the summary of several recommendations from teachers for overcoming difficulties in learning probability.

1. Introduce topics through activities and simulations, not abstractions;
2. Try to arouse in students the feeling that mathematics relates usefully to reality and is not just symbols, rules and conventions;
3. Use visual illustration and emphasise exploratory data methods;
4. Use strategies to improve students' rational number concepts before approaching proportional reasoning;

5. Recognise and confront common errors in students probabilistic thinking;
6. Create situations requiring probabilistic reasoning that correspond to the students' views of the world.

In spite of these recommendations, there still exist some elements of mathematical experience (or approach) integrated to teaching and learning of probability. Behr et al. (1983) and Carpenter (1983) demonstrated that there is a relationship between numeracy skills and students' understanding of probability. One could argue that this is not a strong factor for students failing to understand probability as other factors such as differences in the students such as ability, attitude and intelligence could have influenced their understanding. Rouncefield and Holmes (1989) later developed a practical based approach in teaching probability which intended to provide students with a better feel of the concepts even with some elements of mathematics integrated into it. The issue of practical teaching will be discussed later in this chapter.

In addressing the issue of statistical notations and terminologies, several studies found that students may be able to memorise the formulas and the steps to follow but only seldom appear to get much sense of what the rationale is (Chervany et al., 1977; Garfield, 1981; Kempthorne, 1980). In helping students to select the appropriate formula in hypothesis testing, Soper (1983) used several flow charts to look at the similarities and differences between formulae and in which circumstances they should be used. On the other hand, Ricketts and Berry (1994) investigated the use of a computer-intensive approach in teaching the concepts of probability and sampling distributions and found that it could help students who have difficulty in choosing the right formula to use. Students also sometimes find difficulties in relating samples with the population of interest and why samples are needed to represent the population of interest. Lewis and Brewer (1992) used a 'sample and population' diagram to compare the differences and relationships among the concepts associated with sample and population. One could adopt this approach to suit their teaching needs such as using it as a teaching tool to supplement their usual teaching method. Some introductory textbooks also advocated the use of diagrams in teaching basic concepts. These studies have shown that one can learn and understand statistical concepts without relying too much on the terms and definitions. It can be argued that knowing the definitions does not ensure that the concepts are well understood, certainly not without some form of illustrations and examples.

The issue of students' understanding of descriptive statistics continue with Pollatsek et al. (1981) who discovered that for many students, dealing with the mean is a computational rather than a conceptual act. They found that learning a computational formula is a poor substitute for gaining an understanding of the basic underlying concept and concluded that a student's concept of the mean was usually impoverished and limited to knowledge of a formula rather than of its representational functions. Goodchild (1988) investigated school pupils' (13-14 years) conceptions on the meaning of average as a representative number; as a measure of location and as an expected value. He found that pupils do not reveal an understanding of the full range of the meaning which average is used to convey. The implication is that pupils need to be taught the ideas of representativeness and expectation more fully by working on real data. Kalton (1973) suggests taking examples and set exercises from data relating to the students' substantive subject. One could argue that this is unlikely to be effective if the class consists of students with mixed courses of study. Splitting the students into various classes would not be practical as extra load had to be imposed on teaching staff particularly when there is a shortage of staff.

Loyer (1987) suggested that using classroom data can be useful in illustrating statistical concepts. He discovered that the data given by the class had aroused student interest and had effectively illustrated such concepts as variability, random sampling and the central limit theorem among others. Variability is one of the fundamental ideas, if not the most important concept, in statistics. Students also find it difficult to absorb the notion of variation and the related measures, the variance and the standard deviation, in spite of efforts made to convey the concepts to the students. Several made an attempt at this. Hart (1984) believed that using the mean deviation as an introduction to the standard deviation is appropriate. On the other hand, Wainwright (1984) found it to be more useful to arrive at the standard deviation by starting with a discussion on the estimation of central location parameter. One barrier to the ready acceptance of the notion of variation was thought to be the almost immediate introduction of the formula for the variance of a set of data. Hawkins et al. (1992) noted one easy way to illustrate the concept of variation is by using diagrams or illustrations of data distributions.

By using histograms, Pingel (1993) also illustrated variability based on four different measures of central tendency, i.e., range, interquartile range (IQR), the mean absolute deviation (MAD) and the sample standard deviation (SD). The illustration generalised that the range and to some extent SD and MAD are sensitive to outlying scores while the IQR is not. Following studies on the understanding of the concepts of bar graphs on 121 grade-four (age 8) and 127 grade-six (age 10) students. Pereira-Mendoza and Mellor (1991) found that student had little difficulty reading bar graphs. more difficulty interpreting them, and had major problems knowing when prediction from bar graphs was possible. These studies have shown that illustration of concepts by using diagrams can be helpful as students were able to visualise the outcome of the exercise clearly.

2.3.2 Inferential Statistics

In an introductory statistics courses, about one third of the course content is given over to the general ideas of statistical inference. It is recognised by many educators that this is the section of the introductory statistics syllabus that is most probably prone to misconception and a major cause of confusion (Hawkins et al., 1992; Falk, 1986; Jimenez and Holmes, 1994). Jimenez and Holmes (1994) pointed out the lack of understanding of basic aspects of the overall logic of the procedure for the test of hypotheses. This is partly due to the radically different nature of the type of statistical reasoning for testing hypotheses compared to students' previous experience. On the other hand, Hawkins et al. (1992) quoted that, " students may find difficulty with the classical hypothesis testing approach because their intuitions lead them to seek confirmatory instances, rather than to negate a proposition." These authors mainly pointed out the difficult areas that students have in understanding hypothesis testing and the causes for the misunderstanding. Apart from the fact that these concepts are abstract and difficult to understand, one could investigate ways to make it more understandable and that could possibly minimise student misunderstanding.

Perhaps, because of such difficulties, hypothesis testing is an area in which the *recipe* approach seems to be readily applied. Johnson (1981) recommended teaching hypothesis testing as a six steps process,

STEP 1 Specify Hypothesis.
STEP 2 Decide on Significance Level, α.
STEP 3 Decide on Test and Gather Sample Information
STEP 4 Look up the Critical Value of the Test Statistic
STEP 5 Calculate an Observed Test Statistic
STEP 6 Test and Interpret

Based on the author's previous experience in teaching hypothesis testing to A-Level students in Malaysia, students tend to prefer the six step process as a guideline for testing hypothesis as compared to using the p-value method. Their main weakness, however, is deciding on the appropriate test to use, either the one-sided or two-sided tests and what should be stated in the null and alternative hypotheses.

Some educators perceive the *recipe* approach as a method which fails to help students to learn and understand the concepts of hypothesis testing. One could argue that it would be unethical to completely shut off students who are learning the concept for the first time from this helpful aid. Eventually, students will be shown the slightly harder aspect of testing hypothesis, such as using the p-value method which can be obtained using a statistical package such as Minitab. With these choices, there is a flexibility for students to learn the concept rather than be forced to learn the way they resent most.

There is no doubt that some teachers find it difficult to get students to understand some of the statistical inference concepts. Thus various methods of teaching were developed to help students with the understanding of the concepts. Eckert (1994) showed that a demonstration using playing cards is useful in teaching the basic concept of hypothesis testing which includes using data as evidence against the null hypothesis and determining the strength of the evidence against the null hypothesis, i.e., the p-value. On the other hand, Maxwell (1994) also found that using a coin-flipping exercise to introduce a p-value has helped students in the understanding of the concepts of p-value better. Previously, Rouncefield and Holmes (1989) developed a practical approach to help with the understanding of hypothesis testing and sampling distribution among others. One of the practical sampling exercises require students to take samples of different sizes from a population of blood cholesterol levels of patients. The exercise has demonstrated that the distribution of samples tend to vary in spread and that each sample produced a different range of estimates.

The above studies showed that concepts which are otherwise difficult to understand through the normal lecture method could well be understood through demonstration or practicals. Few studies, if any, have come close to developing a teaching method that could effectively teach students to understand statistical inference concepts without the intervention of aids in conjunction with lectures. This supports the idea of including various teaching methods and materials in the development of the teaching model.

2.4 Practicals and Projects

It is important to include practical statistics as part of the discussion on methods of teaching statistical concepts because many recognised this as an effective method in teaching basic statistical concepts (Rouncefield,1994; Rouncefield and Holmes, 1989; Scheaffer et al., 1995; The Mew Group:exploring statistics; 1993). Its importance is also recognised as there has been an increasing emphasis on doing practicals and projects as part of a statistics course be it in school or at the university level. A review will be made of the necessity of practicals in teaching statistics.

Anderson and Loynes (1987) argued that statistics as a subject is inseparable from its application and practices, and that any statistical education should take account of this connection. Holmes (1994) mentioned that practicals should be used by teachers as a method to introduce or reinforce concepts. Rouncefield (1994) supports this by saying that pupils who have done practicals are actively involved in the learning process and will enjoy and remember what they have done. The drawback is that teachers usually do not have the impetus to do practicals due to time constraints and a lack of appropriate experience in teaching practicals. Some of them find that obstacles arise because of their tight teaching schedules and their obligations to complete a syllabus on time.

There are practical exercises and longer projects. Classroom practicals require less time to conduct and has clear cut objectives as compared to projects. In the former, the teacher has more control and can be geared more easily to specific topics in the syllabus. Projects, which can be considered as an extension to a practicals, have more aims and objectives. Whichever approach was used, careful preparation is needed to maximise students' learning. The importance of using projects are briefly described.

- *Projects put statistics in a context - make them more relevant.* For example, one of the aims of a project is to encourage students to put together their ideas in a context which can be understood by a general reader. The use of appropriate statistical tools and correct computations and interpretations of the analysis are all for a purpose, to make them more meaningful and relevant.
- *Projects give more motivation.* Particularly this is true if students are allowed to choose the subject of the project - something that they really want to know

about. Holmes (1994) reported that a project which was carried out by pupils at Moore County Primary School in an effort to put a village on a gas network has been successful as a result of the analysis of their survey.

- *Projects give a greater feel for real data.* Students who are given the opportunity to collect data on their own would have a better feel when it comes to analysing the data. For example, a group of pupils who were asked to collect data on the height and weight of 20 pupils of age 12 and asked to perform an exploratory data analysis on the data would have a better insights into the practicality of the experiment as compared to those who were just given the data and asked to perform the analysis.
- *Projects emphasise the application of statistics and its usefulness.* In the projects students are forced to draw conclusions and relate them to the practical situation. They also have to think about the best way to report and present their results. This will be easier for students the more projects they do. Furthermore, by interacting with other students doing the same or similar projects they will develop skills of communication and listening about technical issues.
- *Projects show that statistics is not solely mathematics.* In particular the choice of the sample, the design of the experiment or the questionnaire, the practical summary at the end, all require skills that are more than mathematical.

Anderson and Loynes (1987) stated that even though there are arguments from other sources against attempting to teach the skills of practical statistics, this effort is worthwhile and indeed some others argued that the skills of practical statistics are capable of being taught (Committee of Professors of Statistics in the UK and Ireland, Armitage, 1977). The latter argued that the skills of practical may be broken down into simpler abilities. Most crucially, methods for teaching practical statistics have been developed, are in current use, and appear to be effective (Rouncefield, 1994; Scheaffer, 1995; Rouncefield and Holmes, 1989).

2.4.1 Practicals and Computers

In practical statistics, the use of computers is becoming increasingly popular in data analysis. Students need to develop computing skills, in particular the use statistical packages. There are many ways in which computers can be linked with practicals and be used as a tool for practical investigation. For example, spreadsheets have been used to illustrate some of the basic concepts in statistics. Hunt (1995) found that standard deviation can easily be calculated using spreadsheets without needing to repeat the calculation if any of the data values are changed. On the other hand. Holmes (1994) talked about the simulation of experiments where the spreadsheet has been set up to find the binomial probabilites for different values of n and p. thus. by changing the values of p the probabilites change and so do the graphs.

There is a great emphasis on the use of computers as a base for a learning tool. not just from the practical aspects, but to support the learning of introductory statistical concepts. There has been great interest among the statistical educationalist cum computer enthusiast in developing a more interactive computer based learning materials (Marcoulides, 1990; Davies and Radcliff, 1994; Bibby and Davies, 1995; McCloskey et al., 1995). During the past three years, many of these materials have been developed and tested. However, there is still very little extensive research done on the effectiveness of these methods in helping students to understand the statistical concepts taught at the introductory level.

For example, Marcoulides (1990) used an expert system (ES) and a computer-assisted instructional (CAI) programs in an introductory statistics course to assist students in checking their understanding of the steps in the statistical procedures by abstracting from concrete examples rather than by trying to learn abstract procedures and specialise them to specific areas. Davies and Radcliffe (1994) described their experiences in developing computer problem-based teaching and learning materials that allow students to take investigative approach to learning (STEPS project). The problem-based materials provide an interactive learning environment which enable students to explore the practical issues of statistics. Some of the STEPS materials went through several evaluation processes and tested within the class environment at the project sites. Likewise McCloskey et al. (1995) developed a computer based tutorial materials, QUERCUS which aims at teaching basic statistical techniques, report writing and the use of statistics package Minitab for a Biostatistics course. Twelve modules were developed based on the lecture materials and practicals. The modules were tested at the project site using several assessment procedures, for example, the profile questionnaire, the quality assessment questionnaire and the confidence test which aim at investigating how confident were the students in using the computer based tutorial modules. In both the STEPS and QUERCUS projects, the intention is to make the teaching and learning of statistics better, more interesting and relevant. The design and content levels of these materials were impressive.

however more initiatives are needed to investigate their effectiveness in a normal classroom teaching.

In any kind of practical work, the final stage would ultimately require one to interpret the data and provide report for all the work done. This is where the importance of writing as a component to practical statistics should be considered. Radke-Sharpe (1991) highlights that writing helps to assimilate ideas into more thoughtful and meaningful statements. The issue of report writing will not feature largely in this thesis, however, at some point it is required of the students to interpret some of their computational work in class assignments. This is where the skill of even short simple writing demands understanding of the statistical concepts.

2.5 Experiments on Teaching Methods in Statistics: The Methodological Issues

The reviews are made from the perspective of how studies on experimental teaching methods in statistics are conducted, the implications of the design and methods used vis-a-vis the methodology presented in this thesis.

Many studies on experimental teaching methods involved looking at the effectiveness between several teaching methods where treatment versus control group is usually chosen as the design of the experiment (Marcoulides, 1990; Davies, 1996; Keeling and Lenz, 1966; Harding et al., 1981). Marcoulides (1990) looked at the effectiveness of two teaching aids (ES and CAI programs) as measured by students' learning performance in SAT scores. The design involved assigning students at random to three treatment groups with the traditional lecture as the control group based on students' prior achievement levels (high and low) in statistics. A two-factor ANOVA was used to look at the effects of prior achievement and teaching aids and the interactions between them. On the other hand, Davies (1996) determined the effectiveness of two CBL modules in exploratory data analysis where students' performance were measured by a quiz at the end of the learning session. In contrast to Marcoulides' study which assigned students at random using prior achievement levels, Davies assigned students to the two groups (A and B) based on teachers' judgement in matching the ability of the students in both groups. Group A received the two CBL modules while control Group B received paper based practice exercises. One could argue that it is better to use prior achievement level to assign students to treatments rather than based on teachers' judgement as the latter tend to induce bias in the selection process which could possibly be due to human error.

Both studies used post-test scores as a measure of learning performance and prior achievement level as a predictor variable to determine whether the teaching methods used have any effect on the differences in the students. In his analysis, Davies (1996) used regression analysis to predict the test performance of students who used computer based learning materials (CBL) and found that prior attainment in mathematics as a significant predictor for test performance. One could suggest that Davies' study could have also investigated the effects of prior achievement levels, the teaching methods and the interactions between them where prior achievement level could have been used as a covariate in order to adjust the differences between the students in both groups. The fact that these studies have not considered using pretest as a covariate is another factor to be considered in the assessment of a teaching model.

Harding et al. (1981) also matched and assigned students at random to four teaching groups based on prior academic achievement in statistics. The teaching groups were total lecture (L); half lecture and half tape-slides (L-TS); half tape-slides and half lecture (TS-L) and total tape-slides (TS). Their study was designed to see if tape-slide teaching was a reasonable alternative method of teaching to the normal lecture method based on student end of term examination score. Comparatively, Keeling and Lenz (1966) investigated the efficiency of the traditional lecture method (A) with that of programmed textbook approach (B). Students were assigned at random to two groups; Group 1 received method A followed by B and Group 2 received method B followed by B again. This is a modification of a cross-over design. These can be difficult to interpret if there are any carry over or period effects (Senn, 1996).

Compared to Marcoulides' (1990) study, the design in these studies (Harding et al., 1981; Keeling and Lenz; 1966) did not use prior achievement levels to measure the differences between the students apart from using it in matching and balancing the teaching groups. In general, Marcoulides' study design was better than the other studies mentioned here because he does not only take into account the differences between students based on prior achievement level but also the design allowed him to investigate the two factors (i.e., prior achievement and teaching aids) and their interactions.

Several other studies used a different approach in comparing methods of teaching. For example, Borresen (1990) investigated the activities of small group learning in three group conditions; traditional, assigned and voluntary groups. Six introductory statistics classes were used over a period of 3 years, two in each of the three conditions. One of Borresen's analyses involved using analysis of variance to investigate the effects of the group conditions, the semesters and the interactions between them.

In contrast to Marcoulides' (1990) analysis who considered the teaching aids and prior achievement level as predictors to test performance, Borresen considered the factor semesters as possibly having an effect on students' performance. It could be argued that there could be differences between the groups in each semester. As such students' prior knowledge should be considered in taking account the differences between the groups. As it is Borresen made no attempt to control for the ability of the students in the groups. One could argue that there is the tendency that the group conditions may be confounded with the ability of the students in the groups. Keeler and Steinhorst (1994) used the same class over a period of three semesters to compare the two cooperative learning methods with the traditional lecture method. The methods were employed subsequently in between the semesters. They used several approaches but one that allowed students to form pairs by self-selection with roles assigned to group members was found to be successful. This formed the basis of the group arrangement in their study. Self-selection or voluntary association was considered by Keeler and Steinhorst as a strong factor that would increase students' performance. On the other hand, their study also made no attempt to determine or control for the ability of students working in groups which is essential for balancing the groups and in taking account the differences between the students.

In contrast to the pre- and post-tests treatment versus control group design, some studies considered using non-experimental techniques in evaluating methods of teaching. Many articles reported the use of innovative teaching methods as an aid to understanding statistical concepts or as an aid to the already existing teaching methods. In spite of the lack of effort to compare these methods with other methods of teaching, several claimed that their methods were relevant to the teaching of statistics at respective levels (Ricketts and Berry, 1994; Coakley, 1994; Eckert, 1994; Maxwell, 1997). For example, Ricketts and Berry (1994) taught hypothesis testing by resampling amd measure its effectiveness by students' informal response. To some extent they obtained some very positive response but on the other hand, no other measure was used to investigate the effectiveness of the method. In contrast to Ricketts and Berry (1994) who concentrated on just one approach, Coakley (1996) suggested a few approaches to aid the understanding of nonparametric statistics. These include ideas for classroom presentations and projects, inter alia. At each suggestion, Coakley only discussed the merits of each method but did not justify their effectiveness.

One could suggest the use of observational techniques in evaluating the effectiveness of a teaching method in a non-comparative study. Moser and Kalton (1992) indicated the purpose of observation is to probe deeply and to analyse intensively the unit (e.g., a class, a school, inter alia) with a view of establishing generalisations about the wider population to which that unit belongs. Apart from the informal response to the methods used, Ricketts and Berry (1994) could have evaluated students' learning behaviour using Flanders Interaction Analysis system (Flanders, 1970) by keeping track of the non-verbal events that occur while students were using the methods. For example, Adams and Biddles (1970) study on the realities of teaching include making videotape recordings of 32 lessons on 16 classrooms. They claim that this method of observation provides an extremely comprehensive record of classroom behaviour in which the episode can be replayed for the purpose of coding the activity. This enables them to identify and distinguishes the teaching and learning patterns in each classroom. Lambert et al. (1975) also used a variety of techniques in their study including questionnaire and in-depth interview in gathering data from 66 boarding schools throughout England and Wales. One could suggest that the questionnaire and interview techniques would be appropriate for Coakley's (1996) study in evaluating each of the classroom presentations as these would provide a better justification of the effectiveness of the methods used.

2.6 Studies on Teaching Methods in Statistics: The Implication of Results

Several studies indicated the importance of subjects' voluntary participation as opposed to being assigned to treatment groups. The basis of one of Borresen's (1990) hypotheses is that voluntary association may be more conducive to cooperation than involuntary association. However, he found that subjects who were assigned to teaching groups do not perform any better than those in the voluntary group. Keeler and Steinhorst (1995) discovered that students prefer to adjust their groupings in ways which they found natural and which worked for them. They support Borresen's hypothesis and found that students in the assigned roles are more engaged in the course material and learn material better when involved in cooperative groups than when left to work individually. They also found that students who worked in cooperative groups obtained higher mean final scores. Even though they made no comparison between the assigned and the self-selected groupings, there was a positive reaction towards the self-selected group with group structure remaining coherent for long periods of time. These findings also coincide with Dietz's (1993) study that students who were allowed to select their own groups tend to work better in a cooperative manner. The fact these studies have demonstrated that voluntary participation tends to lead to subjects' cooperation as well as to increase test performance is a factor to be considered in the development of the teaching model.

Even though, Dees (1991) found no significant differences between the cooperative and control groups in general , there were some differences between the groups in terms of tests and final examination scores. The groups who were exposed to cooperative methods performed as well or better than the control groups. The findings in the above studies indicate a positive change (even though it was small) in the subjects who were working under the treatment groups. The findings also support the practice of including cooperative work in the development of the teaching model as they indicate that cooperative learning can be successful with effective organisation of the study.

Other studies have involved investigating teaching methods which require little or no intervention from the lecturers. Most tried to investigate the effectiveness of novel methods against the traditional lecture method. For example, Harding et al. (1981) found no significant differences in students' academic achievement or preferences between the lecture and the tape-slides methods. The students perceived the lectures as rewarding and having good presentation, while the tape-slide presentation were seen to be equally rewarding with slightly less agreement of the presentation. Likewise, Keeling and Lenz (1966) found no significant differences between the programme text and the traditional lecture methods. In contrast, Marcoulides (1990) found that the two computer teaching aids (ES and CAI) which were used to help students to correctly choose statistical procedures for data analysis benefit students more than just receiving lectures. In investigating the effectiveness of two computer based learning modules (CBL), Davies (1996) later found that there were no significant differences between the CBL and the paper based practice groups.

The implication of Harding et al. (1981) and Keeling and Lenzs' (1966) studies is that students require some sort of teacher involvement which they can rely on throughout the learning process. One could argue that the value of programmed text and tape-slides depend heavily on the content of the material used and that it is more suitable for teaching skills rather than concepts. The main obstacle to the use of computer teaching aids is the heavy demand on computer facilities, but this would be less problematic in relation to the smaller groups of students. One could suggest that the use of programmed text method, tape slides and computer instructional methods among others with some teachers' intervention would provide a better and more satisfying learning environment for the students. Hence, there is a tendency that teaching can become more effective if teaching methods are combined and organised in a way that would maximise students' learning. Even though research is slowly moving away from the traditional type of teaching into a more innovative type of teaching, the impact of using the methods on students should be considered. One has to be cautious when adopting a new teaching method as failure could damage students' learning.

One other relevant issue is the question of whether or not teaching objectives are clearly defined in determining the content and methodology of teaching. Some educators have researched the development of teaching frameworks and investigated the impact of teaching methods in teaching statistics. Yilmaz (1996) developed a course design which involved teaching statistics using an integrated approach (i.e., computer-intensive approach) and came up with several recommendations for improvement. However, the recommendations are yet to be developed, tested and implemented. Roiter and Petocz (1996), on the other hand, developed a framework for designing introductory statistics courses which are based on theoretical and practical activities but one that takes account of students' anxiety and attitude towards statistics. One could suggest that the latter is an appropriate way of developing a teaching framework because the issue of students' attitude was considered from the outset prior to the development. Although the discussion is brief and does not provide resolutions of all relevant issues, these studies suggest that an effective approach can be developed given the time and resources and one which require effort and cooperation from the teacher and the student in implementing the approach successfully.

2.7 Conclusion

The literature has provided us with a broad view of research into teaching methods but to an extent limited to the issues as regards the development and testing of the teaching model. It should be emphasised that the literature review does not comprise a comprehensive review of teaching methods or teaching statistics but the discussion flows along the lines of the present research objectives. In most cases the discussion is reflected in the choice made leading to the development of the teaching model. The main issues highlighted were teaching methods in statistics, teaching of statistical concepts, the methodological aspects and the implication of the results of studies on teaching methods in statistics.

Several methods of teaching were considered for the teaching model. The lecture or expository method was chosen on the basis that it is still popular in spite of its drawbacks. However, with the intervention of discussion and cooperative work and proper organisation of materials, the combination is believed to work in a normal classroom setting.

In the teaching and learning of statistical concepts, some studies investigated the difficulties that students have in learning basic statistical concepts while others provided solutions from their own teaching experience to make the teaching of statistics better or more understandable. If one could establish a connection between these investigations, it would provide a strong foundation in the development of any teaching model. In Section 2.3, an attempt was made to provide the connection and to take account of the issues developed in the previous studies towards the development and organisation of the teaching methods and materials.

One of the objectives of this thesis is to explore the use of several methodologies of educational experimentation. First is the issue of the design and analysis of the experiment. In general, treatment versus control group design is much preferred in comparing the effectiveness between teaching methods. This is followed by various analyses which are usually reflected in the design and the objective of respective studies. While variations of analyses are possible in the treatment versus control design, some studies extend their analysis to comparing the mean differences between test performances as a result of the teaching methods used and in some cases, using a two-factor analysis of variance procedure to look at the effects of the predictor variables and their interactions. On the other hand, few studies on teaching methods, if any, have come close to using a factorial block design in conducting experimental studies on teaching methods in statistics. This design was also found to be the most suitable for the present study. With this, one could go further and investigate the effects of the teaching methods and their interactions.

Most studies being reviewed have indicated the feasibility of matching and assigning subjects at random to treatment groups. It is recognised that this is important as a way of apportioning out or controlling for extraneous variables. One could argue that forcing students to be part of the study may not truly reflect the results of the study. One could also argue that the effects of the teaching methods does not solely depend on how the subjects are organised in the groups but also how the teaching is organised to maximise learning. Some studies found that voluntary participation is a better alternative as it leads to more cooperation than those being assigned to the groups.

Students' prior achievement levels (or pre-test) in mathematics or statistics were used to assign subjects at random to treatment groups and to make the groups as comparable as possible. Pre-test was used also to measure the differences between the subjects in the groups but is not taken into account in assessing the differences between the subjects in the groups or in measuring the effect of build up learning experience prior to the experiment. In the pre- and post-test control group design, one should be able to test whether the pre-test exercises any significant influence on the performance in the post-test. One should be given an option as follows:

(1) to conclude that the effect is entirely due to the experimental treatment, or

(2) it could have resulted from an increased sensitisation among the experimental subjects as a result of the pre-test.

Thus the effect of build up learning as a result of prior knowledge needs to be controlled. Thus the inclusion of pretest as a covariate to control for initial differences between the subjects in the group is necessary.

A post-test on the other hand, is used in many experimental studies to measure the effect of a teaching method. Whether the post-test is given immediately after treatment intervention or sometime later depends on the objectives of a study. Most studies measure the effects of teaching methods immediately after treatment intervention in order to maximise subjects' participation. One could argue that if the test was exercised much later (e.g., a week or a month later) after the treatment, it is likely that some students would not turn up or would have decided to drop out of the study. If participation is voluntary, a low response rate is even more likely. Thus, the use of post-test immediately after treatment intervention is preferred.

The implication of using the observational technique is that it is suitable for recording the non-verbal events in a classroom. One could argue that in comparing between methods of teaching, observational techniques may not be a suitable approach since the effects of the teaching methods could not be measured solely by observing the behaviour in a classroom. The observational technique would be appropriate if it reflects the objectives of the study and that the use of other methods of data collection could defeat the intended objectives. One serious drawback is the time needed for the observation. To make generalisations on students' learning behaviour would require the observer to intensively record each and every movement of each student over a long period of time. If more than one setting were to be observed, then more than one observer would be needed to record the observations. In the latter, the disadvantage is that differences in the observers are confounded with methods of recording the observation which could lead to distortion of results. Thus by experimenting on students and measuring the effect of each method is a better way of justifying the effectiveness of the teaching methods.

Chapter 3

A FEASIBILITY STUDY ON CLASSROOM OBSERVATION

3.1 Introduction

This chapter serves two purposes. Firstly, we report on the methods used to teach Introductory Statistics at the University of Strathclyde through a small scale investigation based on the use of observational studies. Secondly, it is a feasibility study on the use of observational studies as a methodological technique in the process of developing and testing a statistical education model. Two forms of observational studies are considered: classroom observations of lectures, tutorials and laboratories and a self-report questionnaire of student learning styles.

Chapter 2 described the use of observational techniques by others to investigate teaching behaviour and students' learning activities which are not specific to the teaching of statistics (Lambert et al., 1975; Adams and Biddle, 1970). Hence, as part of the process of developing and testing a teaching model, we investigate whether classroom observation is a feasible and reliable method for looking at the general conduct of classes and the use of different lecturing styles in statistics. As a preliminary exercise, an attempt was made to identify different styles of lecturing using two evaluation instruments.

The second part of the feasibility study is an attempt to investigate students' learning patterns. The methodological technique involved the use of a self-report questionnaire to investigate students' preferences of teaching methods and their preferred learning styles. With this approach we attempt to discover factors which may influence the learning styles adopted by students in statistics and also to see how students think that teaching of statistics can be improved.

This investigation of both types of observational studies is expected to reveal several important findings which would influence future decisions on the appropriate methodological techniques to develop and test the teaching model.

3.2 Classroom Observational Study

The study was conducted at the University of Strathclyde. It began in mid-October 1993 shortly after the beginning of the research and lasted until the end of December 1993. Permission was obtained prior to the observational studies by verbal agreement between the author and some lecturers in the department. The majority of the lecturers who were approached and were told about the observational studies indicated their willingness to cooperate. The courses which were involved in the study were Introduction to Statistics (service course), Probability and Statistics (for second year Mathematics and Statistics), Probability and Statistical Theory, Statistics for Engineering and Statistics for the Pharmaceutical Sciences. There were three types of classroom setting involved in the observational studies, i.e., lecture, tutorial and computer laboratory. Each setting was observed with its own instrument which was developed for the purpose.

In all the settings, the author acted as a non-participant observer. In the lecture setting, the teaching was observed and activities recorded using the teaching skills evaluation form. An attempt was made to identify the lecturing styles using four different forms. In the tutorial setting, the class activites were recorded using the tutorial session observation form and in the laboratory setting, a questionnaire was used in an attempt to investigate students' understanding of statistical concepts.

3.3 Observational Instruments

In this section the feasibility of using self-developed instruments in classroom observations is described. In the lecture setting, a teaching skills evaluation form was adapted from one obtained at a Teaching Methodology Workshop (ITM, 1992) attended by the author. The instrument was developed and used as part of the workshop activity to identify various skills in teaching. It is recognised that this is not the sole method to identify teaching skills and that other observational techniques should be used to support it. This leads to the construction of the second instrument following Brown and Bakhtars' (1988) studies on styles of lecturing. These styles were reviewed in Chapter 2, Section 2.2. This instrument was chosen because of the success reported by Brown and Bakhtar (1988) in identifying different styles of lecturing.

Brown and Bakhtar (1988) used a self-report questionnaire consisting of structured, semi-structured and open-ended questions. They applied factor and cluster analysis to the responses of the 60 structured lecturing styles items and identified five distinctive styles of lecturing representing lectures from the Humanities, Sciences, Biomedical Sciences and Engineering. In the present study, it was not feasible for a single observer to identify all 60 items, and so the researcher selected a representative samples of 30 items and omitted replicated questions [see Appendix A3(ii)]. Additional assistance would have allowed all 60 items to be identified.

Another technique of observing classroom teaching, namely, Flanders Interaction Analysis (Flanders, 1970) could be used. With this instrument the researcher keeps track of selected events that occur during classroom interaction. This technique would be effective if more observers were recruited to record the interactions in the classroom. This is to ensure that every event and movement is recorded. The importance of detailed observation is recognised, however, given the constraint of one observer in observing teaching, this technique was not pursued. Section 3.4 describes briefly how the observations were carried out and in which circumstances the instruments were used during the observation.

3.4 The Report

3.4.1 Lecture Setting

Six lecturers of statistics were observed. Each lecturer was observed on two separate occasions. The first observation involved investigating the lecturers' teaching skills using the teaching skills evaluation form in Table 3.1. The second observation involved identifying and noting down the lecturer's personal teaching styles using the lecture styles forms [see Appendix A3(ii)]. The author did a checklist by noting the presence of the items in the teaching. The more items being noted on the form, the closer it emulates that particular style as indicated on the form.

Table 3.1: Teaching Skills Evaluation Form

VOICE	Clearly audible, well modulated	5	4	3	2	1	Largely inaudible, monotone
PACE	Appropriate number of words and ideas per minute	5	4	3	2	1	Too fast or too slow, t oo many or too few ideas
NON VERBAL COMMUNICATION	Good use of gesture and eye contact, no distracting mannerism	5	4	3	2	1	Poor or distracting use of gesture and eyes, annoying mannerisms

ORGANISATION AND PREPARATION	Well organised eg. introduction to topic, planned repetition, summaries, links between ideas well explained 5 4 3 2 1	Badly organised, confusing presentation
USE OF AUDIO-VISUAL AIDS	Clear, well presented, appropriate use, supportive to presentation 5 4 3 2 1	Cramped, illegible, inappropriate use, detracts from presentation
ATTITUDE	Friendly, responsive, enthusiastic, creates positive impression 5 4 3 2 1	Creates negative impression, dull unfriendly, unresponsive
MONITORING	Aware non-verbal behaviour of students 5 4 3 2 1 (bored, dreaming, puzzled, etc)	Not aware non-verbal behaviour of students

The teaching form is based on a rating scales of 1 to 5. For instance, if the lecturers' voice was clearly audible and well modulated, number 4 or 5 was circled. The process continued down the list with scores added up to produce an average score for the teaching skill. Out of the six lecturers observed, four were thought to have acquired a similar style of lecturing, which is the Exemplary style. One lecturer was categorised under the Amorphous style while the other one showed the Visual style of lecturing. None of them were categorised as Self-Doubters. The majority of the lecturers observed were experienced lecturers, having taught statistics for at least 10 years. The fact that they were exposed to academic teaching for a long time may have accounted for the good teaching skills which they acquired over the years. It should be noted that three of the lecturers taught large groups of students (200 to 300 students) as opposed to the other three who taught smaller classes (30 to 50 students). In order to teach a large class the lecturer has to be well organised with a lot of control over the class. It is difficult to have interactions and the lack of them may be a function of the class size as well as lecturers' preference. One could argue that it is difficult to determine whether large classes would lead to well organised lectures or well organised lecturers were given large classes because they are able to do them successfully. All these could be investigated given the time and appropriate techniques for it.

3.4.2 Tutorial Setting

Tutorial teaching is one of the most common methods used in the teaching of second and third year statistics courses. In view of that, it was decided to look at the conduct of two tutorial classes. The first tutorial session observed was conducted in adjunct to a course in Probability and Statistics for the second year students and the other was conducted in adjunct to a Probability and Statistical Theory course for the third year students. Most of the students took these courses as part of the requirement towards completing their degree. There were several tutorial classes conducted by different tutors and lecturers in these courses. Due to timetabling, the researcher was not able to observe the other classes apart from these two. The tutorial classes were observed on two separate occasions where two experienced lecturers were appointed to conduct the tutorials. The observational outcomes of the tutorial classes are described next.

Case Study 1: Observe the conduct of tutorial session for Probability and Statistics course

There were 15 students in the class. An instrument again adapted from one obtained at the Teaching Methodology Workshop (ITM, 1992) was used to record the general activities of the class during the 50 minutes period. The instrument was used by the author as part of the workshop activity in evaluating teaching. The outcome of the activities recorded in this particular tutorial session can be seen in Table 3.2.

Table 3.2: Observational Outcome of the Tutorial Session (Case Study 1)

INQUIRY ITEM	GENERAL DESCRIPTION
1. HOW IS THE CLASS CONDUCTED? - 'In general.	Tutor generally discuss the previous tutorial solutions. Tutor uses OHP and blackboard to explain materials and solutions. More towards lecturing style-didactic. Tutor did most of the talking - 80% teacher centred
2. ARE THERE ANY SMALL GROUP TASKS? - Describe the task.	None
3. IS THERE ANY SMALL GROUP DISCUSSION? - Describe the situation.	None
4. ARE THERE ANY QUESTION AND ANSWER SESSIONS? - Describe the response from the students. (Active, Neutral, Passive)	After 20 minutes - no question from students. After 23 minutes - two students responded to tutor's questions. After 40 minutes - four students responded to teachers questions. General response from students - rather passive, no initiative to ask questions.

5. ARE THERE ANY BRAINSTORMING SESSIONS? - If so, describe the situation.	None
6. ARE THERE ANY BUZZ GROUPS? - If so, describe the situation.	None
7. ARE THERE ANY INTERACTION BETWEEN THE LECTURER AND STUDENTS? - Describe the situation.	Distant interaction. Interaction came about when tutor began asking questions. Very few responded to the questions.
8. DESCRIBE THE TUTORIAL MATERIALS GIVEN TO THE STUDENTS. - Are they to be handed in? - If so, when?	Tutorial assignments were given 1 week in advance. They are to be handed in and graded. Students are required to hand in 4 questions where 2 questions are discussed during the tutorial.
9. STUDENTS' REACTION IN THE TUTORIAL SESSION? - Describe their attitude (alert, bored, interested, enthusiastic) - Do students ask a lot of questions regarding the tutorial handouts?	Majority are alert but not very responsive. Very few ask questions.
10. OTHER OBSERVABLE FEATURES	Tutor paused several times to check if students understand the materials taught. Tutor is aware of students' reaction.

The tutorial was held in a small auditorium. It was observed that the tutor did most of the talking by discussing the solution to the previous tutorial assignment mainly pointing out the errors that students made in tackling this and similar assignment questions. Students in this group tend to be passive and are less responsive to questions asked by the tutor. There was a distant interaction between the tutor and the students and this came about only when the tutor began asking the group questions. As such, small group activities are rather restrictive. Teaching methods used tend to be didactic or expository. In general, the outcome of the observation tend to reflect the objective and the general (or intended) activity of the class. An informal discussion with the tutor confirmed the above findings.

Case Study 2: Observe the conduct of tutorial session for Probability and Statistical Theory course

There were 34 students with Statistics and Mathematics as their main stream subjects. One tutor was appointed to conduct the tutorial. The same type of instrument as mentioned in Case Study 1 was used to record the general activities of the class. The outcome of the activities recorded in this particular tutorial session can be seen in Table 3.3.

Table 3.3 : Observational Outcome of the Tutorial Session (Case Study 2)

INQUIRY ITEMS	GENERAL DESCRIPTION
1. HOW IS THE CLASS CONDUCTED? - In general.	Tutor allows students to work freely on the tutorial handouts. At times he discusses some of the questions with a group of students. Other times, he walk around the class, generally going to students who ask questions. Most of the time tutor made immediate response to those who asked questions.
2. ARE THERE ANY SMALL GROUP TASKS? - Describe the task.	Generally a big group task. Tutorial exercise is not particularly intended as a small group task.
3. IS THERE ANY SMALL GROUP DISCUSSION? - Describe the situation.	They are not particularly asked to form a small group discussion. However, the students tend to discuss with their partners or someone sitting close to them.
4. ARE THERE ANY QUESTION AND ANSWER SESSIONS? - Describe the response from the students. (Active, Neutral, Passive)	Generally, on a one to one basis and to small groups of 2 to 3 students. Questions are particularly intended for certain groups of students who are already involved in doing certain questions from the tutorial handout.
5. ARE THERE ANY BRAINSTORMING SESSIONS? - If so, describe the situation.	None
6. ARE THERE ANY BUZZ GROUPS? - If so, describe the situation.	Yes

7. ARE THERE ANY INTERACTIONS BETWEEN THE LECTURER AND STUDENTS? - Describe the situation.	Certainly there are more interactions between the students and the tutor and among the students themselves. Because of the way the students are seated make it easier for the tutor to move about allowing him to interact freely with the students.
8. DESCRIBE THE TUTORIAL MATERIALS GIVEN TO THE STUDENTS. - Are they to be handed in? - If so, when?	Tutorial sheets are given out at the beginning of the class. Students are expected to read and try out the questions. They were given 2 weeks to do the problems and expected to discuss them during the tutorial.
9. STUDENTS' REACTION IN THE TUTORIAL SESSION? - Describe their attitude (alert, bored, interesting, enthusiastic) - Do students ask a lot of questions regarding the tutorial handouts?	More than half of the class ask questions with regards to the handouts. 3 of the students are quite attentive and alert. They tend to correct the errors made in the handout. It is observed that the same group of students tend to ask questions. Only a handful seem to be enthusiastic about their work. After half an hour, some of them tend to get bored. Some students do not seem to know how to solve the problems. A few of them read their lecture notes in an attempt to solve the problems. Those sitting at the back of the class tend not to ask questions at all.
10. OTHER OBSERVABLE FEATURES	Tutor tends to create a more friendly learning environment - humourous with the students. Not particularly strict with students' conduct.

The tutorial was held in a classroom where the students were seated in small groups of two to five people. The situation allowed the tutor to interact freely with the students and frequent interactions took place. Even though the tutorial was meant for students to discuss their work, it was noticed that a few students tended to be on their own, reading lecture notes and doing their own personal activities.

Clearly, there were differences in the conduct of the tutorials even though the activites are pointed to the same objective. The activities were restricted by the class setting and the teaching styles of each tutor.

3.4.3 Laboratory Setting

There were two aims to these observations. One was to observe the general conduct of two Introductory Statistics laboratories and two, a small investigation on students' learning of the course lecture using a short questionnaire. Two out of a possible seven labs were observed on separate occasions and the author was an observer in both labs.

There were 16 students in each lab session and two to three tutors were assigned to each lab. Students were given a handout containing the Minitab instructions which was prepared earlier by the course lecturer and were required to perform a task and produce some output. Most students were working with the person sitting next to them while a few others tended to work on their own. The majority of the students observed were able to follow the instructions while a few required more help from the tutors. There were frequent interactions between the students and the tutors.

In investigating students' understanding of the normal distribution concepts which were taught a week prior to the lab sessions, 15 students were asked to fill in a short questionnaire during the last 15 minutes of the lab session. This was a preliminary attempt at an interview study of retention of material previously taught.

Table 3.4: Number of students who responded to the questions

Question 1	Normal distribution	Other than Normal distribution	Don't Know
What was the topic of the lecture	9	4	2
Question 2	Continuous Probability Distribution	Discrete Probability Distribution	Don't Know
Is the Normal distribution a Continuous Probability Distribution or a Discrete Probability Distribution?	9	4	2
Question 3	Bell-Shaped/Symmetrical	Others	
What is the shape of the Normal distribution	13	2	
Question 4	Yes	No	Don't Know
Do you know how to use the Standard Normal Tables to find:			
(a) P(0<Z<1)	13	2	0

(b) P(Z>a)=0.05	4	4	7

From Table 3.4, most students knew the basic concepts of the normal distribution. More than half of the students could recall how to find the probability area of between 0 and 1 without referring to their notes but very few could recall how to find the z-value given the probability area.

During the informal interview with the students, they mentioned of their difficulties in understanding the statistical terms associated with the concepts, the definitions and the formulas as to when and how to use them. They also reported that they sometimes referred to the textbook to learn about Normal distribution and preferred to be given printed notes on the topic rather than having to copy it down.

3.5 Students' Learning Patterns

3.5.1 Introduction

The second part of the feasibility study is an attempt to identify students' learning patterns in a real classroom setting. The reason being that the teaching model to be developed will be tested on some of these students. Thus the choice of suitable components for the teaching model and the process which would influence the implementation of the model is the main focus. This is not a comprehensive study on students' learning patterns. As such the conclusion of the study is only confined within the population mentioned and the learning styles investigated. The implication is that the result may not reflect the learning patterns in other statistics courses. However, this does not invalidate the work as it reflects the aims of the study. The issue is whether the methodological technique used is feasible so that it could be repeated in another study.

The study intended to investigate students' opinion of lecture methods and their learning preferences. One effective way of obtaining these informations from a large sample is by using a self-report questionnaire. The use of direct interviews was considered (and preliminary studies were initiated, see previous section) but would take longer to investigate. Both techniques are subjected to criticism of bias and leading questions (Cohen and Manion, 1989). The advantage is that the number of respondents who can be reached using a questionnaire is extensive and the overall reliability is fair compared to interview (Tuckman, 1972). As such, the self-report questionnaire technique will be pursued. Stephenson (1990) used only questionnaire responses to gauge students' reactions to the usefulness of Minitab in an introductory statistics course. Their conclusion led to the recommendation of use of the statistical package. It can be argued that the use of questionnaire as the sole instrument can lead to important decisions being made. This is what the author hopes to achieve though there are undoubted drawbacks of using just the one methodology. Even when an interview is used to elicit information from the respondent, unexpected things can crop up which were not planned for in a questionnaire study. In the latter, one can only answer questions which were posed while an interview allows one to elicit more information from the respondent. The initial interviews in the labs formed the basis of the questions presented in the study of students' learning which also served as a pilot for the self-report questionnaire.

3.5.2 The Instrument

Lecture-Assessment Questionnaire 1 was given to students at the beginning of the class and returned at the end of the lecture period. A total of 263 students from 380 responded to the questionnaire. The questionnaire has three sections: (A) General, (B) Students' Opinion on the Lecture Methods, and (C) Students' Opinion on their Learning Styles [see Appendix A3(i)]. Section (A) identifies the student's course of study and gender. Section (B) investigates the student's opinion on the lecture methods. This was summarised into several items under student's opinion of the lecture methods in Questions (3) and (5). Section (C) investigates the student's opinion on their styles of learning the subject. Question (6) focused mainly on the way the student understands the materials covered in a lecture. Questions (7) and (8) intend to investigate the student's difficulties in solving the questions from the textbook and what they normally do when they have difficulties in solving the questions. Finally, Question (9) focused on students' difficulties in dealing with Minitab.

The genesis of the self-report questionnaire came from a number of teaching evaluation questionnaires developed by the Center for Academic Practice at the University of Strathclyde. The choice of questions were constructed based on the observation of the teaching on the Introductory Statistics course. The author is aware of the bias if the same questionnaire were to be used to assess teaching in other classes. This is a valid criticism but it is valid for all measurement instrument. It can be argued that the questions were constructed with the aim of investigating one lecturer's teaching method. The questions confirmed the presence of items in the lecturer's teaching but the extent in which each item was emphasised during the teaching from the students' point of view is not known. This is the basis for using Likert scale ranging from Strong Disagree to Strongly Agree to allow for various judgement to be made.

3.6 The Analysis of the Questionnaire

The results of the questionnaire analysis are discussed. Section 3.6.1 describes students' opinion on the lecture methods and how they understand materials covered in a lecture. The latter was investigated based on students' gender and their difficulties in solving questions from the textbook. A chi-square test of independence is used in the analysis. Section 3.6.2 describes students' opinions in different courses of study as regards the lecture methods used and students' learning styles. A chi-square test was also performed to investigate the association between the lecture methods used with students' learning styles. Prior to the test, the validity of performing the chi-square tests was investigated. Whenever necessary, the scales were recoded, by combining categories, to avoid low expected values.

3.6.1 Students' Opinion on the Lecture Methods

This section investigates students' views and opinions on the lecture methods used and the extent to which they think the methods have contributed to their understanding of the subject. This is accomplished by discussing Table 3.5.

Table 3.5: Students' Opinion on Preferred Teaching Methods

	Percentages		
OPINION	Strongly Agree & Agree	Neutral	Disagree & Strongly Disagree
a) Give more lecture notes for copying down. (n=262)	41.2	37.8	21.0
b) Explain lecture materials clearly. (n=263)	79.9	16.7	3.4
c) Give individual attention. (n=262)	14.9	48.1	37.0
d) Pause for questions during the lecture. (n=263)	34.9	43.4	21.7
e) Present content of topic in small units. units. (n=261)	61.7	31.8	6.5

f) Stress or summarise main points during the lecture. (n=263)	91.3	6.8	1.9
g) End lecture with a summary of the topic. (n=263)	85.6	11.4	3.0
h) Make lecture as simple as possible. (n=263)	77.9	16.7	5.4
i) Give detailed printed notes for each lecture. (n=263)	82.1	13.3	4.6
j) Use more examples or case studies to stress concepts. (n=263)	61.2	31.9	6.9

Table 3.5 shows that nearly 80% of the students prefer that lecture materials should be explained clearly, 60% prefer that contents of topics should be presented in small units, over 90% prefer that main points should be stressed or summarised during the lecture, 85% prefer that lecture should be ended with a summary of the topic, 78% prefer that topic lecture should be made as simple as possible, 82% prefer that detailed printed notes should be given for each lecture and 60% prefer that more examples should be used to stress concepts.

Table 3.6: How Students Understand Materials Covered in a Lecture?

	Percentages		
LEARNING STYLES	Strongly Agree & Agree	Neutral	Disagree & Strongly Disagree
a) By listening to lecture only. (n=246)	29.3	22.8	48.0
b) By listening and reading lecture notes. (n=255)	77.3	16.9	5.9
c) By reading the textbook as well. (n=252)	67.1	21.8	11.1

In terms of materials covered in the lecture (n=263), Table 3.6 shows that 48% of the students disagree that they would understand the lecture by listening only. In fact, the typical learning styles showed by majority of the students indicated that they would prefer to listen and read lecture notes with 10% fewer agree that they would read the textbook in order for them to understand the lecture.

Gender

Prior to the chi-square tests, the validity of performing the chi-square tests were investigated and the scales were recoded to avoid low expected values. There were no associations between gender and most of the factors that the students felt would contribute to their understanding of the lecture apart from wanting to be given detailed printed notes (χ^2=12.4, df=3) and prefering lecture materials to be explained clearly (χ^2=11.9, df=3). It was found that 25% of male students as compared to 44% of female students strongly agreed that the lecturer should explain the lecture materials clearly and 51% of male as compared to 31% of female students strongly agreed that they should be given detailed printed notes for the lecture.

With a chi-square value of 13.7 (df=4), a significant association at the 5% level, between gender and students' manner of understanding the materials (i.e., by reading the textbook as well) was noted. A higher percentage of female students (76%) prefer to read the textbook as well in order for them to understand materials covered in the lecture as compared to their male counterparts (60%).

There is also a significant association at the 5% level, between gender and students' difficulties in solving the problems from the textbook (χ^2=21.3, df=4). It was found that 72% of female students (n=99) encountered difficulties in solving problems from the textbook as compared to 41% of their male counterparts (n=118). One reason for the association is that a higher percentage of males did not do the problems from the textbook (31% of males compared to 17% of females). If only those who attempted the questions are considered, there is still a significant association between gender and the difficulites in solving the questions from the textbook (χ^2=15.4, df=3) with females experiencing more difficulties.

Textbooks

Ninety-three percent of the students (n=262) bought the recommended textbook (Introductory Statistics by Weiss & Hassett) for the course. Out of those, 44% thought that the textbook was good. Out of those who owned the textbook and responded to Question 7 (n=227) in the Lecture-Assessment Questionnaire 1, 57% found difficulties in solving the problems from the textbook.

The chi-square test also indicated that there is a significant association between reading the textbook and having difficulties in solving questions from the textbook

(χ^2=19.0, df=4). Out of the students who found difficulties in solving questions from the textbook (n=129), 78% often refer to textbooks. For those who found no difficulties in solving questions from the textbook (n=43), 67% of them referred to the textbook. The reason for the significant association is that out of the students who did not do the questions (n=56), only 46% used the textbook. There were no significant associations between factors which enable the students to understand lecture materials better (Question 5) and the manner in which the students understand the lecture materials (Question 6).

3.6.2 Students' Opinion in Different Courses of Study

Lecture Methods

This section investigates if students with different study backgrounds differ in their opinions about the methods used by lecturers. Concerning what the lecturer should do to enable the students to understand his lectures better (n=263), on average 82% of the students from these courses of study preferred to be given detailed printed notes, 80% wanted the lecture materials to be explained clearly and 78% preferred simple lecture materials. However, there are no major differences among the courses of study in terms of students' preferences of the teaching methods.

On average, more than 50% of the students from each course of study agreed on five out of the six categories on their opinion of the lecture method that lecture materials should be explained clearly, content of topics should be presented in small units, lecture materials should be made as simple as possible, detailed printed notes should be given and more relevant examples and case studies should be used to stress concepts.

Learning Styles

This section investigates if students with different study backgrounds differ in their opinions in terms of their learning styles for this course. Similar learning patterns were noted. The majority of the students from each course of study agreed that they would not understand the materials covered in a lecture by listening to the lecture only. Students majoring in IBML (61%) indicated their agreement on that more than students from the other courses of study. More than 50% of the students from each course of study agreed that they understand the lecture materials from listening, reading lecture notes and by reading the textbook as well. Students majoring in Business Studies indicated their preferences to listening and reading lecture notes (84%) as well as reading textbook (81%). On the other hand, students majoring in Mathematics & Statistics and Accounting understand the materials better by listening and reading lecture notes (85% and 74%, respectively) and less from reading the text (55% and 59%, respectively). Even though the percentage of those in favour of reading the textbook fall slightly, the majority seem to agree that the textbook still plays an important role in helping them to understand the lecture.

The results from this study also show that 40% of all students (n=263) indicated that they sometimes found difficulties in solving questions from the textbook. On the other hand, 16% indicated that they did not find the questions from the textbook difficult while 25% reported that they don't know anything. Nineteen percent did not do the question at all.

In terms of the approach that students in this course would take when they have difficulties in solving questions from the textbook, a high percentage of the students (63.2%) admitted to not seeing the tutor at all whenever they have difficulties in solving questions from the textbook. In fact, the typical styles shown by the students in handling the situation is by discussing the problems with their friends and referring to the lecture notes and texbooks. About a third of the students also admitted that sometimes they did not do anything about it. It was also found that students majoring in Accounting, Business Studies and IBML were more likely to not do anything whenever they have difficulties in solving questions from the textbook.

3.6.3 Summary

As was to be expected, students felt that the lecture can be made more understandable when materials were explained clearly with the main points summarised in the lecture. Students preferred that the contents of the topic should be presented in small units with more illustrations and relevant examples needed to stress the concepts. They also agreed that this should be made as simple as possible. They felt that detailed printed notes would further help them in understanding what was taught during the lecture. The majority of the students mentioned that they understood the lecture materials from listening and from reading lecture notes and textbooks. They felt that the level of their understanding would improve if enhanced by their preferred lecture methods. They also felt more comfortable discussing their work problems with their friends, at the same time referring to the lecture notes and textbooks.

3.7 Conclusion and Discussion

Direct Observations

The study has provided information on the use of methodological techniques which are feasible in the context of statistical education and other techniques which could be used to improve methods of classroom observations. The classroom observations were intended to be small scale for the purpose of providing the background information for future studies.

It is difficult to investigate teaching styles and students' learning without giving much consideration to the instruments used. It was the author's choice to observe teaching styles in statistics using the Brown and Bakhtar's (1988) technique in identifying styles of lecturing. They use a more rigorous approach but not specific to the teaching of statistics. The author is aware that the instrument as developed for the statistical teaching environment has lower reliability due to the omission of several items. Even though an effort was made to choose items which identify the different styles of lecturing, it would have increased reliability to have used all 60 items. As mentioned earlier, more than one observer would be needed and certainly more time would be required for the observations.

Indeed there is bias in recording the items with the author as the observer and this could be reduced by recruiting more observers which would also provide information on the reliability of the scoring system. The drawback of using more than one observer is that the differences between the observers are confounded with the evaluation methods. Being a non-participant observer should not have any effect on the behaviour of the subjects being observed in the lecture. On the other hand, there is more likely to be an effect in the tutorials and laboratory as the observer appears more visible.

The methodological techniques used have some implications. To identify the lecturing styles would definitely require a thorough investigation using questionnaire and interview techniques such those of Brown and Bakhtar's studies (1989). A more sophisticated technique such as Flanders Interaction Analysis system (Flanders, 1970) and other techniques such as videotaping, questionnaire and in-depth interview (Adams and Biddle, 1970; Lambert et al., 1970) could supplement the techniques used in the present study to further analyse teaching behaviour and interactions. The study could be improved by giving more consideration to the time, subjects and the type of instruments used in recording the observation. Repeated visits to the classrooms should assist the use of instruments in recognising and identifying teaching behaviour but this was not the aim of this study.

The main thrust of this thesis is to develop and investigate the feasibility of testing a teaching model in a real classroom setting. One of the reasons for the feasibility study is to see whether it is worthwhile pursuing direct observations further. There are flaws in the methodological techniques used, especially when only six lecturers, two tutorials and two laboratory sessions were observed. As such, no conclusive decision can be made on this apart from reporting occurrences using the available instruments. The weakness with this type of investigation is that the effectiveness of a teaching method could not be measured due to the nature of direct observational study which is based on natural occurences or events.

In view of the findings of direct observational study, it was felt that the teaching model cannot be developed and tested by solely observing teaching regardless of what instruments are used. This is the main reason for not pursuing them any further.

Self-Report Questionnaire

In the investigation of students' learning patterns, the implications of the methods used is that it is feasible to study those associated with one lecturer's style of teaching with the instrument specially developed for the purpose of improving his or her own teaching. It can be argued that the questionnaire was constructed following the aims of the investigation and to extend it to other investigations would require some revisions of the questionnaire.

The disadvantages of using the questionnaire is the inability to ask more probing questions and the poor rate of return (Tuckman, 1972). This is a valid criticism but since the investigation was conducted in a normal lecture setting, the response rate was quite high (263 out of 380) because those who attended the lecture were asked to fill in the questionnaire. Due to the feasibility of using the questionnaire technique, this will continue to be used in future investigations because of its ease in obtaining information in a classroom.

The difficulty was felt in structuring the questions that would ensure a high rate of response. This is a valid criticism for any newly developed questionnaire. The bias is recognised when the questions were constructed based on the observations of the teaching of one senior lecturer in an introductory statistics course. With hindsight, it is hard to see why anyone should not agree in answer to some of the questions on lecture preference. It would have been possible to ask respondents to rank the items in order of importance. But this does not invalidate the work as it reflects the aim of the study. There is a scope for the self-report questionnaire study to be implemented in a different class. One could suggest an improvement to the structure of the questionnaire, possibly a different phrasing of the questions to get it closer to the objective of the study.

The study of students' learning patterns was considered successful because some of the findings were important for the development of the teaching model. The main conclusion is that majority of the students prefer the passive learning styles rather than the active learning styles. This in some ways reflect how the subject was taught, that is through an expository method. This method is predominantly the way of teaching even after much criticism on its effectiveness. This can be rectified given good organisation, clarity of explanations and interesting presentation (Gage and Berliner, 1979; Brown and Amstrong, 1984). In view of this, the expository method is taken into account in the development of the teaching model.

Printed notes and worked examples were also used in the teaching. The findings from the study indicated that students prefer the notes and worked examples. However, in a large group teaching there were almost no interactions between the teacher and the students and there is bound to be a group of students who may not benefit from the work in the lecture. At this point, it is felt that interactive learning would bring more benefits to disadvantaged students. Dietz (1993) and Cummings (1983) found that learning together involved positive interdependence and promoted interaction among students. Even though performance varies from student to another, Webb (1982) found that better performance was associated with active participation in the group. So the issue here is to try and develop a method which is going to ensure that there is a greater cooperative and interactive work. Therefore the idea of students working cooperatively will be considered. In general, concept teaching through worked examples, notes and cooperative learning will form the basis of the components of the teaching model because it is felt that there is a scope for these components to help improve the teaching of statistics.

Chapter 4

LARGE SCALE EXPERIMENT

4.1 Introduction

One aim in Chapter 3 was the investigation of the appropriate methodological techniques which can be used in pursuing classroom investigations. The 'observe and record' type of observational technique was not feasible for investigating the *effectiveness* of a teaching model as opposed to observing its conduct, and thus is not pursued further. As part of the process of developing a teaching model, other methodological techniques are now investigated.

An experiment comparing two methods of delivering lecture materials in a normal classroom setting is described here. This coincides with several studies where methods of teaching were compared in several experimental teaching groups (Harding et al., 1981; Keeling and Lenz's, 1966) in a classroom setting. In the present study, students' understanding of statistical concepts were compared after using different techniques of presenting lecture contents in the Introductory Statistics class. The method of delivery is mainly expository. The aims of this study are to see whether:

(1) an experiment can be conducted in a large classroom setting,
(2) the experiment can be well controlled in such an environment,
(3) it is possible to influence what students retain about concepts by manipulating the lecture presentation,
(4) it is possible to use experimental techniques in statistical education successfully, and
(5) it will be feasible to conduct a similar experiment in the future.

With reference to aim (3) above, the purpose is to find out to what extent students are able to grasp and understand a lecture immediately after the lecture is delivered. This leads to the question, can methods of delivery influence students' understanding of the lecture? This is accomplished by investigating student responses on Lecture-Assessment Questionnaire 2 [see Appendix A4].

4.2 Design of the Experiment

4.2.1 Experimental Procedure

It was decided to conduct the experiment with first year undergraduate students taking the Introduction to Statistics course during the spring semester of 1994. There were two classes attending the same lectures at different times, one on Monday and the other on Wednesday, and these formed the experimental groups. After some preparations, the experiment finally took place in mid-February 1994. Prior to the experiment, students had had lectures on Sampling Distributions and Confidence Intervals. The next topic was Significance Testing, covering the ideas of hypothesis testing using the critical value and p-value approach. There were two reasons for the choice. One, this topic came at the convenient time of the study. Two, significance testing is one of the key concepts that pose a number of problems to students. This was highlighted in the literature review in Chapter 2. One study pointed out that the lack of understanding is in the overall logic of the procedure for hypothesis testing (Jimenez and Holmes; 1994). Hawkis et al. (1992) also discuss the difficulties that students have in understanding hypothesis testing. The importance of researching into this area is recognised. This led to the investigation of how manipulation of the contents of the lecture (critical value and p-value) might have an effect on student understanding of hypothesis testing. The experimental lecture involved showing simultaneously the testing procedures for the critical value and p-value approach. Most textbooks adopt the critical value approach whereas computer programs such as Minitab tend to provide information on p-values. For those who rely heavily on the textbook in teaching the concept, it is assumed that the critical value approach is preferred because of the impression that it is easily understood by many students.

It was decided that the lecturer should deliver the lecture to the groups in two different ways. Every step was taken to ensure that the experiment was conducted accordingly. Prior to the lecture, students in both groups were briefed about the experiment and what was required of them. They were given a handout listing the objectives of the lecture, the key concepts of the topic and what was expected of them at the end of the lecture. A set of lecture printed notes was also given. The idea was to enable students to concentrate on the lecture by listening and referring to the printed notes rather than listening and copying notes at the same time.

4.2.2 Teaching Methods

It was decided that for the Monday group, the concept of the p-value approach rather than the critical value approach would be stressed. The lecturer started off by referring to the list of key concepts of the topic and explained to the class what he expected the students to know at the end of the lecture. Following that was a demonstration, the Pepsi-Cola tasting challenge. The idea was to get some of the students involved in a classroom activity, at the same time providing the group with a general idea of the concepts of hypothesis testing. These include the null and alternative hypothesis, p-value, critical value, critical region, test statistics and significance level. The lecture continued by first showing the p-value approach through a worked example. The lecturer also used other examples quoting articles from newspapers. This was followed by the critical value approach which was briefly discussed by referring to the lecture printed notes. The fact that the p-value method of hypothesis was stressed first led to the decision to test the hypothesis that students who are taught more fully using the p-value approach respond with more correct answers on p-value questions. Finally, the lecturer illustrated and briefly explained the circumstances of the Type I and Type II errors by giving notes for copying down.

For the Wednesday group , the lecturer initially started off with the list of key concepts and what he expected the students to know at the end of the lecture. He then proceeded by showing a worked example using the critical value approach followed by the Pepsi-Cola tasting demonstration. This time the demonstration was done after the lecturer had introduced the critical value approach first. Following the demonstration was a brief discussion on the p-value approach. Examples using articles from newspapers were also used to illustrate the importance of understanding the concept of significance testing. The hypothesis was that students who are taught more fully using the critical value approach are expected to respond with more correct answers on the critical value questions. The lecture closed with an explanation of the Type I and Type II errors where students were given notes for copying down. At the end of both lectures, students were required to fill in a questionnaire designed to assess their understanding of the lecture materials.

There may be bias as only one person is involved in teaching the materials. The bias is recognised from the point that the lecturer will be more familiar with the materials on Wednesday than on Monday. However, it is argued that the effect will be very small since the materials were delivered according to how they was prepared with emphasis given to the respective key concepts. Another aspect of the experiment design is consideration of two different lecturers with each lecturer teaching each group of students. This type of design would cause serious disruption to students' learning as students normally had the same lecturer. Also the differences between the lecturers would be confounded with the styles of teaching and the ability of lecturers which would lead to problems with the interpretation. Thus the use of one lecturer is the best choice.

4.2.3 The Instrument

A self-developed questionnaire [see 'Lecture-Assessment Questionnaire 2' in Appendix A4] was the main instrument used to assess student understanding of the lecture. The questionnaire is divided into four sections: (A) General - Course of Study and Gender. These two attributes will be used later in the analysis to determine whether there are gender differences and course differences with regard to those answering the questions. (B) Key Concepts - To list as many key concepts as they can remember being mentioned in the lecture. This is to determine if they list the concepts which the lecturer believes to be important and which were stressed during the lecture. (C) Class Materials, and (D) Lecture and Comments. In (C), there are nine True-False Questions 4 to 12. Questions 4, 5, 6 and 11 were related to the p-value approach while Questions 7 to 10 were related to the critical value approach. The only multiple choice Question 12 was related to the Type I error. Question 4 was to see if students know the definition of p-value. Question 5 was on the application of p-value and decision making. Question 6 was to find out if students know the decision rule for a hypothesis testing using p-value.

Questions 4, 5, 6 and 11 will determine whether students attending the Monday's lecture in which case are exposed to the p-value concept first manage to answer those questions correctly. Questions 7, 8, 9 and 10 will determine whether students attending the Wednesday's lecture, in which case are exposed to the critical-value concept, manage to answer those questions correctly. Questions 7 and 8 were used to test students' ability and knowledge in making a correct decision in testing hypothesis using the critical value. Question 9 was to find out if students are able to locate the critical region when a test statistic is compared to the critical value. Question 10 was used to test whether students are able to make the right conclusion in a hypothesis test or otherwise. Question 11 was to test if students are able to make a decision on the hypothesis test if the p-value is not given. Finally, Q12 was to test students' ability in defining Type I error. The reason for using conceptual and not computational questions is that the latter would more likely to provide information on computational skills rather than conceptual understanding. In (D) - Students' perception of the lecture quality plus comments and suggestions about the class, course and lecturer.

The difficulties in choosing the appropriate questions to ask are recognised. The choice of questions should inform whether or not students understand statistical knowledge and provide information on student learning (Garfield, 1994). So the questions were constructed with that purpose in mind. However, with only nine true/false conceptual questions the information needed may not be sufficient. If more questions were to be asked, a different time would have to be allocated for the assessment. As this assessment is made on the day of lecture, the amount of lecture materials would have to be reduced, thus defeating the purpose of assessing learning. If multiple choice questions were to be used, more time would be required to administer the test. So given the circumstances of the study condition, the methodological techniques have made the best use of limited resources available.

4.3 Descriptive Analysis

4.3.1 Univariate Analysis: Lecture Groups, Gender, Course of Study, Perception

About 380 students were registered for the course with 110 students attending on Mondays and 270 on Wednesdays. They were students from various courses of study such as Business Studies, International Business and Modern Languages (IBML), Mathematics and Statistics, Computer Science, Science Studies and Arts and Social Sciences. Only 96 students from the Monday group and 153 from the Wednesday group responded to the questionnaire. There were also more males (54.2%) than females (45.8%). Students majoring in Business Studies contributed 42% of the sample followed by Mathematics and Statistics which is 20%. This is shown in Table 4.1.

Table 4.1: The Percentages of Students in each Course of Study

	Course of Study					
	Business Studies (n=105)	IBML (n=40)	Maths/Statistics (n=50)	Computer Science (n=40)	Science (n=10)	Others (n=4)
Percentage of sudents	42.1	16.1	20.1	16.1	4.0	1.6

Overall perception of the course as indicated in Table 4.2 shows that the majority of the students rated the lecture as being fair to good.

Table 4.2: The Percentages of Students' Perception on the Lectures

	Perception				
	Very Poor (n=17)	Poor (n=35)	Fair (n=117)	Good (n=59)	Very Good (n=17)
Percentage of Students	6.94	14.3	47.8	24.1	6.9

4.3.2 Bivariate Analysis

In this section, the analysis of association between two categorical variables is described. A chi-square test analysis shows no significant association between gender and lecture groups (χ^2=3.39, df=1). There is a significant association, at the 5% level, between course of study and lecture groups (χ^2=74.6, df=4). One of the differences is that of the IBML students, 5% went to the lecture on Wednesday and 95% went on Monday (see Table 4.3) . The differences are also attributed to the Business Studies and Mathematics/Statistics groups in which more than 75% of the students went to the lecture on Wednesday.

Table 4.3: Students in each Course of Study with regard to Lecture Groups

	Row (Percentages)	
Course of Study	Monday (n=96)	Wednesday (n=153)
Business Studies (n=105)	23.8	76.2
IBML (n=40)	95.0	5.0
Maths/Stats (n=50)	22.0	78.0
Computer Science (n=40)	40.0	60.0
Science Studies (n=10)	20.0	80.0

There is no association between lecture groups and students' perception of the lecture (χ^2=3.44, df=4). At the 5% level, there is a significant association between gender and course of study (χ^2=30.4, df=4). Table 4.4 shows that a greater proportion of male students are found to be majoring in courses such as Mathematics and Statistics, Computer Science and Science Studies. On the other hand, 75% and 51.4% of female students can be seen majoring in courses such as IBML and Business Studies, respectively.

Table 4.4: The Male and Female Students in each Course of Study

(Row) Percentages

Course of Study	Male (n=135)	Female (n=114)
Business Studies (n=105)	48.6	51.4
IBML (n=40)	25.0	75.0
Mathematics/Statistics (n=50)	62.0	38.0
Computer Science (n=40)	85.0	15.0
Science Studies (n=10)	70.0	30.0

Chi-square test analysis (χ^2=32.6, df=4) show a significant association, at the 5% level, between gender and perception. Of the male students 82.5% thought the lecture was at least fair. The corresponding percentage for the females students was 74.3% (see Table 4.5). Finally, there was no association, at the 5% level, between perception and course of study (χ^2=22.8, df=16).

Table 4.5: Students' Perception of the Lectures with regard to Gender

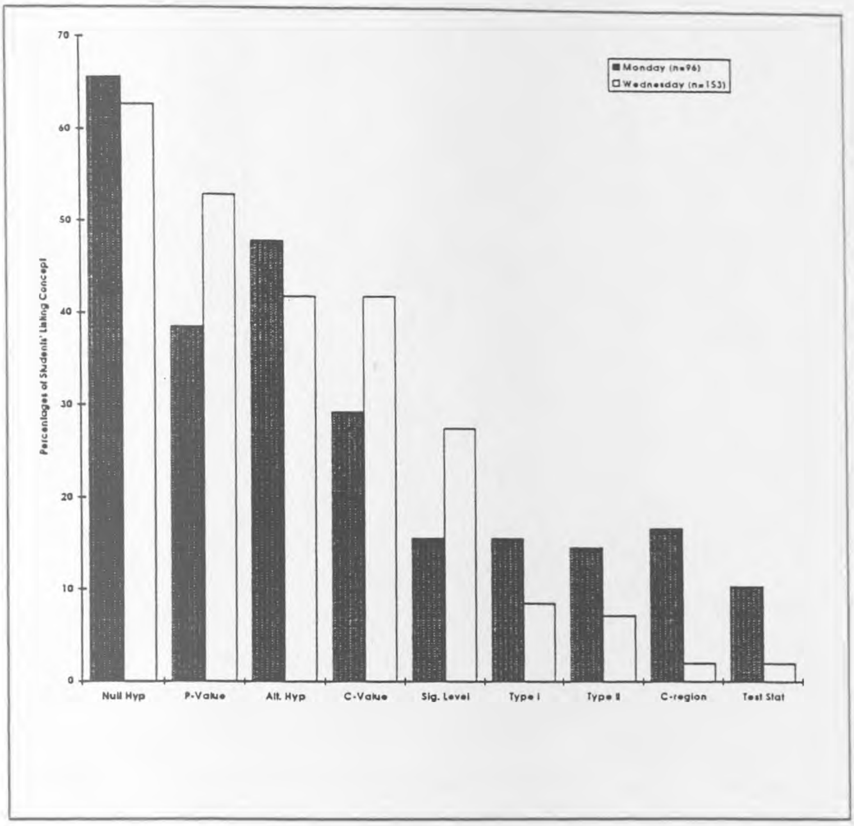
(Row) Percentages

Gender	Very Poor	Poor	Fair	Good	Very Good
Male (n=132)	3.0	14.4	43.9	29.5	9.1
Female (n=113)	11.5	14.2	52.2	17.7	4.4

4.3.3 Analysis of Key Concepts

This section will look at the distribution of the number of concepts listed by students with regard to lecture groups, gender and course of study. Figure 4.1 shows that the percentage of students listing Null and Alternative Hypothesis as key concepts are more or less comparable between the two lecture groups.

Figure 4.1: **Bar Chart showing the Percentages of Students' Response to Key Concepts in each Lecture Group**



Comparing between students' listing P-Value and Critical Value as the key concepts, it was found that slightly more students from the Wednesday's group responded to these two key concepts than students in the Monday's group. However, no significant differences were detected between the lecture groups with regard to students listing of key concepts. The percentage of students who listed Significance Level, Type I error, Type II error, Critical Region and Test Statistics as key concepts are relatively small in both groups.

The percentage of male students listing key concepts are largely comparable to the female students. Overall the percentage of students listing the key concepts are quite low; for most concepts the percentage is below 50%. Also, no differences were detected between the course of study with regard to students listing of key concepts.

4.3.4 Analysis of Correct Answers

The percentages of students' answering Questions 4 to 12 correctly with regard to lecture groups, gender, course of study and perception are described in this section. The results are presented in the form of two-way contingency tables. Wherever possible, chi-square tests of independence are used in association to the tables to investigate for any significant differences between the attributes.

It is expected that students on Monday will respond with more correct answers on Questions 4, 5, 6 and 11 than on Wednesday. For Questions 7 to 10, it is expected that those in the Wednesday group will respond with more correct answers than on Monday. For Question 12 (i.e. types of error) equal proportion from each group are expected to answer this question correctly because the concept was taught briefly to both group of students. The analysis proceeds by discussing the percentage of students' answering correctly in the lecture groups, gender and course of study.

By Lecture Groups

The percentages of students' answering the questions correctly are compared between the lecture groups to find out if students who were shown the concept of p-value approach first and more intensively, respond with a higher percentage of correct answers on the p-value questions. The percentage of those not answering correctly included those students who did not attempt the question. This is accomplished by discussing Table 4.6.

Table 4.6: Percentages of Students' Answering Correctly by Lecture Groups

	MONDAY (n=96)		WEDNESDAY (n = 153)	
Question No.	% of students' answering	% of students' answering correctly	% of students' answering	% of students' answering correctly
Q4 (P-value)	97.9	61.5	96.7	76.5
Q5 (P-value)	97.9	74.0	96.1	55.5
Q6 (P-value)	100.0	71.9	97.4	64.1
Q7 (C-value)	97.9	47.9	95.4	43.8
Q8 (C-value)	98.9	58.3	95.4	60.8

Q9 (C-value)	96.9	45.8	94.8	44.4
Q10 (C-value)	97.9	34.4	94.1	46.4
Q11 (P-value)	98.9	59.4	94.8	50.3
Q12 (H-test)	95.8	53.1	92.8	71.9

Key:

Test is significant at the 5% level

Table 4.6 focuses on the percentage of students' answering the questions (Q4 to Q12) and the percentage of students' answering the questions correctly. Comparatively, the percentages of students' answering each question are slightly higher on Monday than on Wednesday. On Monday, out of 96 students, 97.9% answered Question 4 and 61.5% answered the question correctly. On the other hand, out of 153 students on Wednesday, 96.7% answered Question 4 and 76.5% answered the question correctly. For Questions 7, 9 and 10 the proportion of those answering correctly is less than 50% of the total of those answering the questions.

The results show that there were significant differences in the percentage of correct answers between the lecture groups with regard to Questions 4, 5, 10 and 12. A chi-square test analysis (χ^2 =7.69, df=1) shows that there was a relationship between students' answers to the question on the definition of p-value (i.e. Question 4) and lecture groups. In spite of the fact that the p-value approach was taught first and stressed much more during the Monday's lecture, the percentage of those answering correctly is higher on Wednesday than on Monday.

A chi-square test analysis (χ^2=7.87, df=1) also indicated that there is a relationship between students' answers to this p-value question (Question 5) and lecture groups. It was found that 74% of the Monday group answer this question correctly while 55.5% from the Wednesday's lecture group did so. For Question 6, the intention is to find out if students know the decision rule for a hypothesis test using p-value. Even though, the percentage of those answering correctly on Wednesday was found to be slightly higher than on Monday, however, the chi-square test analysis between students' answers to Question 6 and the lecture groups was not significant. The test also shows no significant difference, at the 5% level, between students' answers to Questions 7, 8 and 9 and the lecture groups. In Question 10 where students were tested on their ability to make the right conclusion on the hypothesis test using the

critical value approach, chi-square test analysis (χ^2= 4.66, df =1) between students' answers to Question 10 and lecture groups was found to be significant, at the 5% level. With regard to Question 10, it can be seen that on Monday, 34.4% answered the question correctly while on Wednesday, 46.4% answered the question correctly.
In Question 12, students are tested on their ability to define the Type I error. Chi-square test analysis (χ^2 = 12.6, df = 1) shows that there is a significant difference, at the 5% level, between students' answers to Question 12 and lecture groups. The percentage of students answering Question 12 correctly on Wednesday is 71.9% while the corresponding percentage on Monday is 53.1%. Even though the percentage of students' answering the question on Wednesday is slightly less than on Monday, a greater percentage of students from the Wednesday's group responded to the question correctly. This may suggest that students in the Wednesday group understood the concept on types of error better than on Monday.

By Gender, Course of Study and Perceptions

The percentages of students' answering correctly with regard to gender, course of study and perceptions are described. The intentions are to investigate: (1) which gender group responded with a greater percentage of correct answers in each question, (2) if students in different courses of study responded with different percentages of correct answers in each question, and (3) if students who responded with greater percentage of correct answers have good perception of the lecture.

The percentage of female students answering is consistently slightly lower than the percentage of male students answering. In all cases, except for Question 4, there were no significant differences between students' answering the questions correctly with regard to gender. In Question 4, the percentage of male students (77.8%) answering correctly is higher than female (62.3%).

Table 4.7 focuses on the percentage of students' answering correctly in each course of study. Emphasis will be placed on the four main courses of study, that is, Business Studies, IBML, Mathematics & Statistics and Computer Science students taking the Introductory Statistics. The reason being that they are the only large groups and represent about 80% of the total class population. Due to small number of students repesenting the Science Studies and Other courses, these two were omitted from the analysis. The chi-square test analysis shows a significant difference, at the 5% level, between students' answers to Question 4 (χ^2 = 11.0, df = 3) and course of study. Table 4.7 shows that the percentage of Computer Science students (82.5%) responded with more correct answers in the question about the definition of p-value (i.e. Question 4) as compared to other courses of study. No significant differences were found between students' answering correctly and the course of study in the other questions.

Table 4.7: Percentages of Students' Answering Correctly by Course of Study

Questions	Business Studies (n=105)		IBML (n=40)	
	% of students' answering	% of students' answering correctly	% of students' answering	% of students' answering correctly
Q4 (P-value)	96.2	74.3	95.0	52.6
Q5 (P-value)	95.2	66.0	97.5	71.8
Q6 (P-value)	97.1	76.5	100.0	67.5
Q7 (C-value)	95.2	47.0	95.0	60.5
Q8 (C-value)	95.2	60.0	97.5	64.1
Q9 (C-value)	94.3	45.4	92.5	43.2
Q10 (C-value)	93.3	46.9	97.5	30.8
Q11 (P-value)	94.3	55.5	97.5	64.1
Q12 (Htest)	93.3	71.4	92.5	51.4

Questions	Mathematics/Statistics (n=50)		Computer Science (n=40)	
	% of students' answering	% of students' answering correctly	% of students' answering	% of students' answering correctly
Q4 (P-value)	98.0	79.6	100.0	82.5
Q5 (P-value)	96.0	56.3	100.0	67.5
Q6 (P-value)	98.0	63.3	100.0	60.0
Q7 (C-value)	96.0	39.6	100.0	45.0
Q8 (C-value)	96.0	56.3	100.0	72.5
Q9 (C-value)	96.0	60.4	100.0	37.5
Q10 (C-value)	96.0	50.0	97.5	33.3

Q11 (P-value)	96.0	64.6	100.0	45.0
Q12 (Htest)	94.0	72.3	97.5	76.9

Key: Test is significant at the 5% level.

Table 4.8 indicates that apart from Question 11, the test for all questions show no significant differences between students' answering correctly and their perception of the lecture. The chi-square test analysis ($\chi^2 = 10.3$, df = 2) shows a significant difference, at the 5% level, between students' answers to Question 11 and their perception of the lecture. This indicates that the more they liked the lecture, the greater is the percentage of correct answers for this question.

Table 4.8: Percentages of Students' Answering Correctly with regard to their Perception of the Lecture

	Poor (n=52)		Fair (n=117)		Good (n=76)	
Questions	% of students' answering	% of students' answering correctly	% of students' answering	% of students' answering correctly	% of students' answering	% of students' answering correctly
Q4 (Pval)	94.2	79.6	96.6	66.4	100.0	77.6
Q5 (Pval)	92.3	66.7	97.4	63.2	98.7	65.3
Q6 (Pval)	96.2	68.0	98.3	73.0	100.0	61.8
Q7 (Cval)	92.3	47.9	97.4	45.6	98.7	49.3
Q8 (Cval)	94.2	65.3	97.4	58.8	98.7	66.7
Q9 (Cval)	94.2	36.7	98.3	48.2	98.7	52.0
Q10 (Cval)	94.2	42.9	94.9	41.4	100.0	47.4
Q11 (Pval)	94.2	46.9	97.4	49.1	98.7	70.7
Q12 (Htst)	94.2	63.3	94.9	71.2	93.4	69.0

Key:

Test is significant at 5% level

4.4 Conclusion

Some of the results appeared to be in favour of the research hypotheses and some were not. In terms of students' answering correctly, those who attended the Monday lecture and were exposed more thoroughly to the concept of the p-value responded with a greater percentage of correct answers on three out of four questions on p-value (Q5, Q6 and Q11). Those on Wednesday responded with a greater percentage of correct answers on two out of four questions on critical values (Q8 and Q10). Even though there was a significant difference between the groups in Question 5, there is not enough evidence to conclude that those students who attended the Monday's lecture and were more thoroughly exposed to the p-value concept, responded with a greater percentage of correct answers on the p-value questions. Similarly for Question 10, even though there was a significant difference between the groups, there is still not enough evidence to conclude that those students who attended the Wednesday's lecture, who were more thoroughly exposed to the critical value concept, responded with a greater percentage of correct answers on the questions. The experiment has shown that there is only a marginal effect in terms of understanding the concepts between the two lecture groups.

In the case of those answering correctly with regard to gender groups, there is insufficient evidence to conclude that one gender group is better than the other. The percentages of those answering correctly are more or less comparable apart from Question 4. In terms of course of study, the evidence that students with Mathematics and Statistics and Computer Science major should perform better than students from the Business Studies major is not conclusive. In the case of those listing key concepts with regard to lecture and gender groups, the experiment failed to conclude that one lecture group is better than the other and that one gender group is better than the other. Finally, those who perceive the lecture as good are generally those who have responded with more correct answers to Question 11 only.

One general point is the problem of multiple testing and it is necessary not to go overboard with the interpretation of the marginally significant chi-square test. A total of forty-eight chi-square tests were performed. On balance, there was evidence from the analysis that the lecture groups were an important determinant of correct response but these are possibly confounding effects with gender, course of study and students' perception of the lecture. To investigate further, a more sophisticated analysis using a Logistic Regression Analysis will be used to look at the differences between the lecture groups adjusting for the effects of gender, course of study and students' perception.

4.5 Analysis of Lecture-Assessment Questionnaire 2 using Logistic Regression Analysis

4.5.1 Introduction

In this section, the effect of four explanatory variables, that is gender, lecture group, course of study and perception of lecture, on the number of correct answers produced by students are discussed. The relationships among four explanatory variables were investigated in Section 4.3.1 as a prelude to this modelling. The number of correct answers is the response variable which is discrete and dichotomous and it is the effects of the remaining variables on this that are of primary interest.

The purpose of this section is to find out whether there exist group differences in terms of correct answers produced by the students (i.e. the response variable) after taking into account gender, course of study and perception of the lecture. At some stages, the researcher will investigate the odds-ratio of answering Questions 4 to 12 correctly with respect to the explanatory variables, which include three categorical variables [group, gender and course of study] and one quantitative (or ordinal) variable [students' perception of the lecture]. For example in the case of gender, the odds-ratio is used to measure how much more likely or unlikely it is for the outcome to be present among the males than among the females. For the quantitative variable perception, 4 students who did not respond to the question on the perception of the lecture were excluded from the analysis. The linear approximation inherent to the use of perception as a quantitative variable was also tested and justified as being valid. This process involved a linear Logistic Regression Model as shown in Figure 4.2.

Figure 4.2: Logistic Regression Model for the Proportion of Correct Answers

$$\boxed{\log\left(\frac{p}{1-p}\right) = \mu + \alpha_i + \beta_j + \pi_k + \gamma_l + e_{ijkl}}$$

where $p =$ proportion of correct answers;

$log\left(\frac{p}{1-p}\right)$ is referred to as a logistic transformation of the proportion p

(or the logistic link function of p);

μ is the overall mean;

α_i is the effect of group i where $i = 1$ *(Monday group)*,

$i = 2$ *(Wednesday group)*;

β_j is the effect of gender j where $j = 1$ *(Male)*,

$j = 2$ *(Female)*;

π_k is the effect of course k where $k = 1$ *(Business Studies)*,

$k = 2$ *(International Business and Modern Languages, IBML)*,

$k = 3$ *(Mathematics and Statistics)*,

$k = 4$ *(Computer Science)*,

$k = 5$ *(Science Studies)*,

$k = 6$ *(Others)*;

γ_l is the effect of perception l where $l = 1$ *(Very Poor)*,

$l = 2$ *(Poor)*,

$l = 3$ *(Fair)*,

$l = 4$ *(Good)*,

$l = 5$ *(Very Good)*;

e_{ijkl} is the residual error term.

Logistic link is used in association with the probability distribution of the observations which in this case, is the Binomial distribution. The formation of linear models in terms of choice of error distributions and appropriate link functions was suggested by Nelder and Wedderburn (1972). One feature of this modelling process is to assess how many parameters are required in the model to provide an adequate description of the data. The aim is to find the model which represents the data adequately. For this purpose, Nelder and Wedderburn suggest a goodness-of-fit criterion based on the maximised value of the log-likehood, known as the deviance. For contingency table data, the deviance is much like the Pearson chi-squared statistic as portrayed earlier in the analysis. In this study, the analysis particularly deals with contingency tables and models to represent the data with a minimum number of parameters.

4.5.2 The Analysis and Results

The analysis in this section discusses the unconditional or unadjusted effect of the independent variable 'lecture-group' on Questions 4 to 12 when it is compared with the conditional or adjusted effect of the independent variable 'lecture-group' in the presence of the other explanatory variables (i.e., gender, course of study and perception). The following summary response variables are also used.

(i) Questions 4 to 12 respectively,

(ii) The total p-value questions (Questions 4, 5, 6, 11) and the total c-value questions (Questions 7, 8, 9, 10).

First, the adequacy of the linearity assumption for perception is tested. The dependent variable is the total number of correct answers.

Fitting $log\left(\frac{p}{1-p}\right) = \mu + \beta_1\,(Group) + \beta_2\,(Gender) + \beta_3\,(Course) + \varphi X(Perception)$

where $\varphi X\,(Perception)$ is a 5-dimensional vector

yields a deviation of 371.48 on 236 df. Constraining perception to have a linear effect results in a deviance of 378.79 on 240 df with the model

$log\left(\frac{p}{1-p}\right) = \mu + \alpha\,(Group) + \beta\,(Gender) + \pi\,(Course) + \gamma\,(Perception)$

The difference in the deviance is 7.31 on 4 df which is not significant and so there is no evidence to suggest that the linearity assumption is invalid.

In checking the Binomial assumption for the number of correct answers, a normal plot of Pearson's residual was constructed. This is shown in Figure 4.3. The graph shows that the residual values lie approximately on a straight line which indicates that the Binomial assumption for the number of correct answers is valid.

Figure 4.3: Normal Plot of Pearson's Residual for the Number of Correct Answers

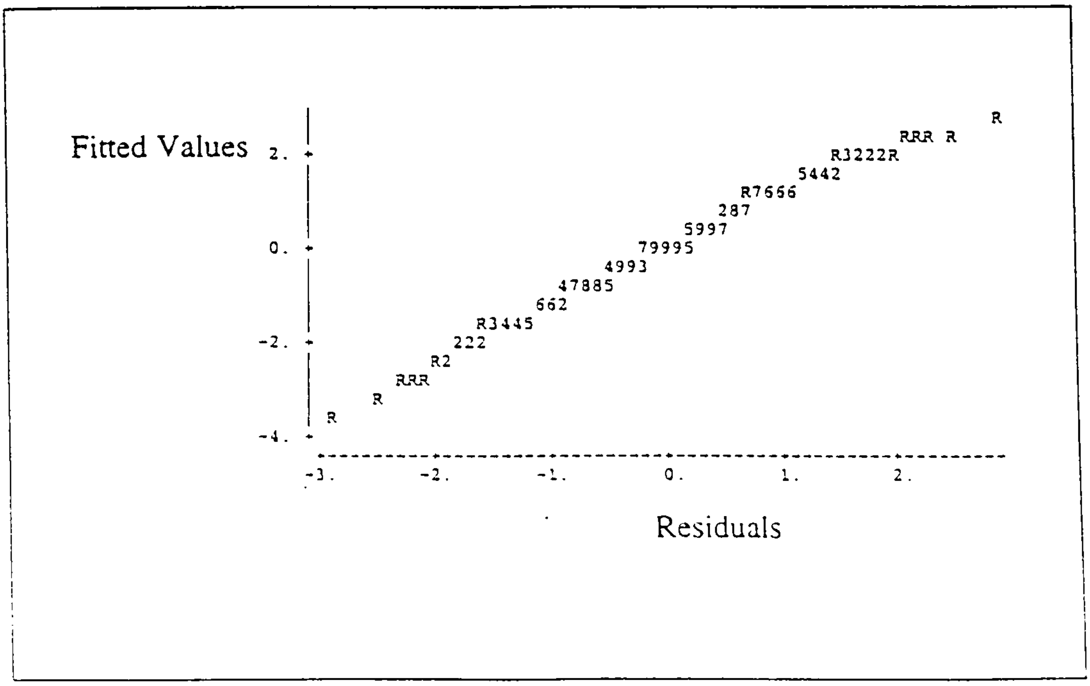


Table 4.9(a) and 4.9(b) show the p-values which are obtained for both the unconditional and conditional effect of lecture-group for each category, respectively. P-values are used in this case to identify the significant difference in the variable 'lecture-group' where a p-value of less than 0.05 are considered to be significant. The intention is also to investigate the effect of the variable 'lecture-group' being insignificant after taking into account other explanatory variables.

Table 4.9(a): Chi-square and p-values for both unconditional and conditional effect of Lecture Group for Questions 4 to 12

	Fit Group only		Fit Group/ (Sex, Course, Perception)	
Questions	χ^2	p-value	χ^2	p-value
Q4	6.3	0.0121	0.6	0.4386
Q5	8.8	0.0047	8.4	0.0047
Q6	1.7	0.1923	1.6	0.2059
Q7	0.4	0.5271	0.94	0.3428
Q8	0.1	0.7518	2.5	0.1138
Q9	0.1	0.7518	0.83	0.3711
Q10	3.54	0.0614	2.84	0.0943
Q11	2.0	0.1573	0.7	0.4028
Q12	9.0	0.0000	5.2	0.0226

Key: p-value is significant at less than 0.05

Table 4.9(b): Chi-square and p-values for both unconditional and conditional effect of Lecture Group for Total p-value and Total c-value questions

	Fit Group only		Fit Group/ (Sex, Course, Perception)	
Category of Questions	χ^2	p-value	χ^2	p-value
Total p-value	2.6	0.1069	4.6	0.032
Total c-value	0.48	0.4795	2.6	0.1069

Key: p-value is significant at less than 0.05

In all of these cases, a backward selection procedure was used to determine the final Logistic Regression Model. The results of these analysis are displayed in Table 4.10 and Table 4.11.

Table 4.9(a) shows that for the unconditional effect of group, the p-values indicate that the lecture-group is significant in Questions 4, 5 and 12. For the conditional effect of group, the p-values indicate that lecture-group is significant in Questions 5 and 12. In Question 4, the reason for lecture-group being insignificant after taking into account other explanatory variables is because of students' course of study. During the backward selection process in Table 4.10, course of study was found to be the only significant variable for Question 4.

Previously in Table 4.6 in Section 4.3.4, 61.5% of students from the Monday's lecture answered the p-value question correctly while a slightly higher percentage, that is, 76.5% from the Wednesday's group answered correctly. It was also found that 38% of the IBML students attended the Monday's lecture and only 2% went on Wednesday. Out of those who went to the Monday's lecture, 47% answered correctly and 53% did not. Part of the difference is found to be explained by student course of study, that is, more IBML students went to the Monday lecture and they answered poorly on Question 4. Student prior ability is a likely confounding effect which has possibly contributed to the effect of no difference between the teaching methods. This confounding effect will be taken into account in the main experiment in Chapter 6. A suitable experimental design will be used to tackle the situation.

Table 4.9(b) shows that when lecture-group was fitted in the presence of other explanatory variables for the category of total p-value questions, the p-value indicated a significance difference between the lecture-groups. This indicates that answering all the p-value questions correctly are not affected by course of study. After fitting the Logistic Regression Model in Figure 4.2 for Questions 4 to 12 using a backward selection process, the final Logistic Regression Model was obtained for each category of questions. The final Logistic Regression Model for Questions 4 to 12 and the odds-ratio of the significant variables are displayed in Table 4.10. Refer to Figure 4.2 for details of levels of factors in the model.

Table 4.10: Final Logistic Regression Model for Questions 4 to 12 and the Odds-Ratio of the Significant Variables

Independent y-variable	Final Logistic Regression Model	Odds-Ratio
Question 4	* $log\left(\frac{p}{1-p}\right)$ = .9163 - .9163(Crse2) + (.215) (.383) .3494(Crse3) + .6343(Crse4) + .4700(Crse5) (.404) (.468) (.819)	B.Studies = 1 IBML = 0.4 Math/Stat = 1.42 C.Science = 1.88 Others = 1.60
Question 5	$log\left(\frac{p}{1-p}\right)$ = 1.044 -.8207(Wednesday) (.232) (.283)	Monday = 1 Wednesday = 0.44
Question 6	* $log\left(\frac{p}{1-p}\right)$ = 1.061-.3300(Crse2)- (.223) (.404) .5714(Crse3) -.6554(Crse4) -1.908(Crse5) (.366) (.392) (.723)	B.Studies = 1 IBML = .72 Math/Stat =.56 C.Science = .52 Others = .15
Question 7	* $log\left(\frac{p}{1-p}\right)$ = -.2103+.5126(Crse2)- (.196) (.375) .2793(Crse3) +.009625(Crse4) -1.176(Crse5) (.351) (.374) (.813)	B.Studies = 1 IBML = .81 Math/Stat = 1.67 C.Science = 0.76 Others = 1.01
Question 8	* $log\left(\frac{p}{1-p}\right)$ = .2877+.2231(Crse2)- (.197) (.382) .1273(Crse3) +.6817(Crse4) -.6931(Crse5) (.345) (.405) (.675)	B.Studies = 1 IBML = 1.25 Math/Stat = .88 C.Science = 1.98 Others = 0.5

Question 9	$log\left(\frac{p}{1-p}\right) = -.9960 + .2593(\text{Perception})$ (.413) (.128)	V. Poor = 1 Poor = 1.30 Neutral = 1.68 Good = 2.17 V. Good = 2.81
Question 10	$*\log\left(\frac{p}{1-p}\right) = .2489 - .5984(\text{Crse2}) +$ (.197) (.397) $.1689(\text{Crse3}) - .4820(\text{Crse4}) + .2489(\text{Crse5})$ (.345) (.391) (.662)	B.Studies = 1 IBML = 0.55 Math/Stat = 1.18 C.Science = 0.62 Others = 1.28
Question 11	$log\left(\frac{p}{1-p}\right) = -.7184 + .2866(\text{Perception})$ (.406) (.127)	V. Poor = 1 Poor = 1.33 Neutral = 1.77 Good = 2.37 V. Good = 3.15
Question 12	$log\left(\frac{p}{1-p}\right) = .1252 + .8141(\text{Wednesday})$ (.204) (.272)	Monday = 1 Wednesday = 2.26

Key: Crse(2)-IBML, Crse(3)-Math/Stat, Crse(4)-Computer Science, Crse(5)-Others;
Values in the brackets ( ) are the standard errors of the parameters of the proportion of correct answers.

* Weighted Model

From Table 4.10, the final logistic regression models for Questions 4, 6, 7, 8 and 10 are found to have been attributed to students' course of study. There were small numbers of students in the Science Studies answering these questions and the proportion of students from this course of study answering correctly is high. Hence, weighted logistic regression models are used to eliminate the effect of that particular course of study.

The odds-ratio of 0.44 (with a 95% C.I. of between 0.253 and 0.766) indicates that students from the Wednesday group are less likely to answer Question 5 correctly than students from the Monday group whereas in Question 12, students from the Wednesday's group are over two times more likely (with a 95% C.I. of between 1.32 and 3.85) to answer correctly than students from Monday's group. Analysis of Question 5 implies that students from the Monday group who were thoroughly shown the p-value concept are more likely to produce correct answers for this question. For Questions 4, 6, 7, 8 and 10, the odds-ratio among students' course of study varies from one question to another. In Questions 9 and 11, the final model shows that perception of lecture has an effect on student answers to the questions. The odds-ratio indicates that the more students perceived the lecture as good, the more likely they were to produce correct answers for these questions.

For the category of Total p-value and Total c-value questions, using a backward selection process, the final Logistic Regression Model were obtained. The final models and the odds-ratio of the significant variables are displayed in Table 4.11.

Table 4.11: Final Logistic Regression Model for the Total p-value and Total c-value Questions and the Odds-Ratio of the Significant Variables

Independent y-variable	Final Logistic Regression Model	Odds-Ratio
Total p-value	$log\left(\frac{p}{1-p}\right) = 0.6931 - 0.2205\ (\text{Wednesday})$ (.108) (.137)	Monday = 1 Wednesday = 0.80
Total c-value	$log\left(\frac{p}{1-p}\right) = -0.5210 + 0.1444\ (\text{Perception})$ (.200) (.0621)	V. Poor = 1 Poor = 1.15 Neutral = 1.33 Good = 1.54 V. Good = 1.78

Key: Values in the brackets ( ) are the standard errors of the parameters of the proportion of correct answers.

The odds-ratio indicates that students from the Wednesday group are 20% less likely to answer questions on p-value correctly than students from the Monday group. For the category of questions on c-value, the significant difference is found to be attributed to students perception of the lecture. The odds-ratio is higher for students who perceived the lecture as good (54%) than for students who perceived the lecture as poor. This indicates that perception of the lecture has an effect on student answers to the c-value questions.

For the total number of key concepts, there was no upper bound to the number of key concepts which could be listed and so, a Poisson log-linear model is used to model the number of key concepts listed. The Poisson log-linear model is

$$\log(\lambda) = \mu + \beta_1\,(Group) + \beta_2\,(Gender) + \beta_3\,(Course) + \beta_4\,(Perception)$$

where λ is the mean number of key concepts listed.

In the final Poisson regression model as shown in Table 4.12, there was no significant effect of group. On average females listed 21% more concepts than the males (with a 95% C.I of between 1.06 and 1.56). Also it was found that the ratios varies among the course of study and that students' performance increases with perception where the mean number of concepts listed for students who perceived the lecture as good are 1.33 times the means for those who perceived the lecture as very poor.

Table 4.12: Final Poisson Regression Model for the Total Number of Key Concepts Listed and the Odds-Ratio of the Significant Variables

Independent y-variable	Final Logistic Regression Model	Ratios
Total number of key concepts listed	$\log(\lambda) = \underset{(.149)}{.504} + \underset{(.0850)}{.190}(\text{Female}) - \underset{(.120)}{.0240}(\text{Crse2}) + \underset{(.103)}{.231}(\text{Crse3}) + \underset{(.122)}{.0856}(\text{Crse4}) - \underset{(.223)}{.102}(\text{Crse5}) + \underset{(.238)}{.638}(\text{Crse6}) + \underset{(.0396)}{0.0938}(\text{Perception})$	Male = 1 Female = 1.21 B.Studies = 1 IBML = 0.98 Math/Stat = 1.26 C.Science = 1.09 Science = 0.90 V. Poor = 1 Poor = 1.10 Neutral = 1.21 Good = 1.33 V. Good = 1.46

Key: Crse(2)-IBML, Crse(3)-Math/Stat, Crse(4)-Computer Science, Crse(5)-Others; Values in the brackets ( ) are the standard errors of the parameters of the mean of key concepts listed (λ).

4.6 Conclusion

Adjusting for the confounding effects, a logistic regression model was used to take into account students' gender, course of study and perception. The test results indicate that these variables are important at some stages. Course of study seems to have an effect on students' performance. This was seen from the results in Questions 4, 6, 7, 8 and 10 which are largely explained by course of study. There is some evidence that the Monday group answered better on the p-value questions even after taking into account gender, course of study and perception. However, the evidence is weak and is not always detected. One possible reason is that there is so much variability over the subjects (i.e. the students). Student educational background may have some influence on their perception of lectures. Hence, it would be desirable to consider educational background into consideration. The result has provided some indications that differences between students are greater than the effects of the manipulation of the lecture and these differences need to be carefully controlled. One way would be to induce a bigger change in the effects of the teaching method. However, this was not feasible in the prevailing circumstances.

4.7 Discussion

The discussions focus on the choice of study methods and the results of the experiment. The implications of these in reference to the development of the teaching model are described.

In Dees' (1991) study, random allocation of subjects to treatment (or teaching groups) was possible as her experiments were conducted in a special (or non-natural) setting. One could argue that doing experiment on teaching methods in a non-natural setting does not present a true picture of a classroom activity and that, in a natural setting, randomisation might disrupt the normal teaching process. The present study was conducted in a natural setting and it was not possible to allocate the students at random to the teaching groups for timetable reasons. It is recognised that a non-random allocation of subjects tends to ignore the effect of extraneous variables. To adjust for this non-random allocation of subjects, the course of study was taken into account in modelling the number of correct answers. Apart from timetable reasons, the decision to use all students in the lecture groups was to increase the rate of participation. If the experiment was conducted outside the timetable framework, there is a possibility that many students will not turn up for the experiment as it is difficult to get the students to come voluntarily to the class. In spite of this effort, not every student was keen to take part in the study. This information was gathered from Part (D) of the questionnaire where some said that they were not prepared to answer the questionnaire as they were told about the experiment only on the day of the lecture.

One drawback in the present design is that students' prior knowledge in statistics was not controlled meaning that there was no pretest to measure prior attainment. The implication for not doing the pretest is that it is not known if the changes in posttest score between the Monday and Wednesday lecture groups are a function of learning due to teaching strategy or a function of the pretest score. Hence, the question of how much a student learns is not known. Likewise Dees' (1991) study used several pretest scores as a diagnostic measure of prior learning but no further check was made to indicate whether the changes in the posttest was a function of the pretest score. On the other hand, Marcoulides (1990) used prior achievement levels to investigate whether the differences in the students have any significant effect in test performance regardless of the teaching methods used. The use of pretest both as a measure of prior knowledge and in measuring the differences between the students is an important area which should be considered in later experiments.

The implication of time was considered. On the test material level, only nine questions could be fitted into the 10 minutes slot at the end of the lecture. A better way would be to fit thirty questions with more choices of p-value and critical value questions but this was not feasible as more time is needed to administer the test within the lecture. Assuming all students came at the same time, it would take about 15 minutes to monitor the pretest at the beginning of the lecture and 15 minutes for the posttest at the end of lecture. The amount of lecture would have to be reduced to allow for assessment to take place. As the emphasis is on the manipulation of critical value and p-value approach, it is important to have these concepts covered adequately in the lecture.

On Wednesday, the c-value approach was shown first while the p-value approach was shown toward the end of the lecture with the Pepsi-Cola task in the middle. On the other hand, in the Monday lecture, the Pepsi-Cola task was shown at the beginning of the lecture, followed by the p-value approach and finally the c-value approach toward the end of the lecture. Table 4.6 showed that 76.5% of the students in the Wednesday group answered Q4 (p-value) correctly as compared to 61.5% on Monday. This indicates that the students in the Wednesday group remembered the p-value concept better since it was shown toward the end of the lecture. The same explanation could be given for Q5 (p-value) in which 74% of the students in the

Monday group answered the question correctly as compared to 55.5% on Wednesday. This also indicates that the students in the Monday group remembered the p-value concept better since it was shown at the beginning of the lecture. There is a possible effect in the order of the materials delivered. In relation to this, McLeish (1968a) also found that material presented at the beginning and end of talk was better remembered than the central sections of the material.

In general, those who were exposed more thoroughly to the p-value approach responded with a higher percentage of correct answers than those who were shown the critical value approach. The implication of the result is that it is important to consider teaching using the p-value approach first followed by the critical value approach. Some researchers emphasised the need for p-value in research analysis and in helping students develop an understanding of significance testing. The importance of using p-value in teaching hypothesis testing was demonstrated by Eckert (1994) and Maxwell (1994) through classroom practical activities. However, they make no comparison with the other teaching methods which they claim were unsuccessful.

The experiment could be improved with a better design. Indeed, the experiment could not show that one method of teaching is decisively better than the other in increasing students' learning. As there was no measure of prior knowledge, posttest score cannot be compared against this to measure learning. This issue of pre- and post-tests are dealt with in the design of small class experiment in Chapter 5. As a result of the experiment, the following conclusions are made.

- It is difficult to conduct a large classroom experiment in a normal setting without seriously disadvantaging the students.
- It is difficult to control an experiment in large classes unless a complex experimental design is adopted in a special setting.
- It is not possible to influence what students retain about concepts by manipulating the lecture presentation due to large variation in course of study.
- Appropriate experimental techniques should reflect the necessary conditions of the study.
- It is not feasible to conduct similar experiments unless changes are made in the design and techniques used in the experiments.

Chapter 5

SMALL CLASSROOM EXPERIMENT: DESIGN AND METHODS

5.1 Introduction

In the literature review of Chapter 2, it was discovered that in most studies the choice of experimental design and methodological techniques used in educational experiments depended upon the objective of the study, the feasibility of using the techniques and resources available to support those studies. For example, Harding et al. (1981) and Keeling and Lenz (1966) chose to perform experimental studies on teaching methods which they have used in their courses. The authority and resources available to them made it possible to conduct such studies. In Chapters 3 and 4, it was discovered that to conduct experiments on teaching methods successfully would require a careful experimental design. The studies in these chapters have highlighted the difficulties in conducting experimental studies in the development of a teaching model due to the use of inadequate methodological techniques and factors which are outwith the control of the experimenter.

Nevertheless, to complement the methodological approach in previous chapters, an attempt was made to design a better experiment for investigating the teaching model involving a relatively complex method for educational testing, i.e., a 2^3 randomised factorial block design with confounding. This is the first time such a design has been used in experimenting on methods of teaching statistical concepts. The design was used to estimate the effects of the components of the teaching model developed in this chapter.

The proposed teaching model has three components—Provision of Notes, Working in Pairs and teaching Concepts through Worked Examples. Issues related to effective methods of teaching are reflected in the choice of these components. Success in cooperative learning (Webb, 1982; Dees, 1991; Garfield, 1993) led to the choice of working in pairs, while notes and worked examples came from the analysis of students' learning patterns in Chapter 3. Teaching statistical concepts is the most crucial aspect in the development of the teaching model. The literature is filled with studies on how to teach statistical concepts effectively but very few have indicated the importance of using diagrams and flowcharts as a tool in teaching basic statistical concepts. Lewis and Brewer (1992) described a simple, yet effective way to introduce concepts using a 'sample and population' diagram to supplement the explanations required in teaching statistical concepts. The use of diagrams or illustrations was also noted by Hawkins et al. (1992) as an easy way of conveying difficult concepts to students. Hence, in the teaching model concepts were taught in worked examples with the aid of a flowchart. Issues related to the effective presentation of lectures (Brown and Bakhtar, 1988; Gage and Berliner, 1979; Brown and Amstrong, 1984) reflect the choice of organisational methods suitable for a number of teaching styles, in essence the Exemplary style (Brown and Bakhtar, 1988).

In the next section, the subjects who took part in the experiment are described followed by the issues raised by the pilot experiment. The main experiment is in Section 5.5 with details of the experimental design in Section 5.6.

5.2 Subjects

The experiment was conducted on the first year students who attended the remedial class for the Introductory Statistics course during the academic year of 1994/95. At the beginning of their first semester, about 400 students who went to the first lecture were asked to provide particulars about their course of study, their highest level of achievement in Higher Maths and whether they had taken any Statistics courses prior to entering the university. Those who had obtained a 'C' or lower in the Higher Maths examination or equivalent were strongly encouraged to attend the remedial classes. In the final selection, 42 such students were identified. The lecturer in charge of the course wrote a letter encouraging them to attend the remedial classes and stressing the importance of attending the classes——that it would be beneficial for them and also because they were among the students who needed help the most. There were timetabling difficulties but an attempt was made to find a suitable time so that as many of the students as possible could participate in the study. Those who finally came were volunteers who wished to get extra help in their studies. To reduce the Hawthorne effect, students were told about the experiment in general but not when the experiment would actually take place.

5.3 The Pilot Experiment

5.3.1 Introduction

Prior to the main experiment, a pilot study was conducted in order to gain experience with the experimental procedures. Specifically, the reasons are to investigate:

(1) the feasibility of the study design and the experimental methods, and

(2) the adequacy of the experimental materials.

To investigate (1), two teaching models were used in teaching the concepts of variability and the normal distribution. These topics were chosen because they are fundamental concepts of statistics and therefore it is natural to pilot on them. Many students also find the variability concept difficult to understand in spite the fact that several ways of illustrating the concepts were used (Pingel, 1993; Hawkins et al., 1992; Hart, 1984; Wainwright, 1984). The normal distribution forms the basis of most statistical inference.

The pilot had to be conducted outside normal teaching hours using the small remedial classes. This was done because of unsuccessful attempts with large classroom experiments in terms of its feasiblility and as it is difficult to exercise control in large groups. One of the difficulties was to decide on a suitable time for the students to attend the classes. This was done by sending out letters encouraging them to come at the time suggested by the lecturer for the course. Teaching materials were developed and their usefulness investigated in conjunction with the teaching model. The intentions were to find out: (a) the amount of material that could be taught, (b) the appropriate assessment to measure students' learning, and (c) the amount of time that should be allocated for each activity in the teaching model. This was done to produce a good design and methodological techniques for the main experiment.

5.3.2 Teaching Methods

Because of the constraint in timetabling and the preparation needed to conduct the pilot, it was decided that only two teaching models can be piloted. The teaching models each contain two teaching components (see Table 5.1). Teaching Model 1 emphasised giving printed notes to the students and working in pairs and Teaching Model 2 emphasised on students working individually with no printed notes given.

Table 5.1: Components of Teaching Models 1 and 2

Teaching Model 1	*Teaching Model 2*
Printed Notes	No Printed Notes
Work in Pairs	Work Individually

The following describes the sequence in which the teaching was carried out during the experiment.

Teaching Model 1:

1. Explain briefly the key concepts to students stressing the importance of understanding the concepts while highlighting the important concepts due for discussion.
2. Go through one worked example which discussed both the concepts and the numerical aspects of the question. The example was chosen to show the proper way of tackling the question which includes the used of the appropriate formula in the calculation. During the explanation process, key-concepts were gradually induced into the worked problem where students were constantly reminded of the key-concepts at various stages of the example. The aim of using this technique was to help student to identify the concepts and learn to relate the concepts in the worked examples.
3. Give out problem exercises for students to solve. These were later collected and marked. The problems chosen were related to the worked examples and key concepts discussed in the class. This group of students were also given an additional set of reading notes designed to help them with the exercise. They were also encouraged to work either in pairs or in groups of three students if necessary.

Teaching Model 2:

In this teaching model, the sequence in which the teaching was implemented is similar to Teaching Model 1 as shown in Steps 1 to 3 above. The difference is that students were not supplied with any printed notes and also they were asked to work individually as far as possible. However, they were encouraged by the researcher to ask questions when they encountered with any difficulties in solving the problems.

5.3.3 Procedures

The experiment was conducted in two consecutive weeks. In each week, students were required to go to either one of the two remedial groups, either on Wednesday or Thursday. Since students' participation were voluntary and their attendances to the classes depend on the time that suits them, it was not feasible to assign them at random to the groups. Variability concepts were taught in week 1 and normal distribution in week 2. Those who went on Wednesday were exposed to Teaching Model 1 and those on Thursday to Teaching Model 2. In week 1, six students came on Wednesday and seven on Thursday while in week 2, nine students went on Wednesday and six on Thursday. Tests were administered before and after the employment of the teaching models. The materials used in the experiment include teaching notes, printed notes on related concepts, handouts on key concepts and definitions, problem exercises and assessment materials (i.e. pre- and post-tests). The following describe the aims of the experiment materials.

Pre-test and Post-tests

- To measure students' learning before and after the employment of teaching model. This includes their ability to recall and use certain formulas appropriately.

Teaching Materials (notes and worked examples)

- To define the key concepts.
- To teach concepts and computation through worked examples using flowcharts and printed notes.
- To help students to identify key concepts and relate it to the worked examples.

Class assignments

- To monitor students' learning and understanding through problem solving.
- To look at the extent to which students can provide answers adequately.
- Training session in problem solving.

5.4 Conclusion

It was discovered that students choose to attend the classes which suits them best. Students' participation was very low even though much effort had been put through to encourage students to come to the class. In terms of teaching materials, too much had to be taught within a short period and the same test materials were used to assess students' learning before and after the teaching. The bias was recognised as practice effects tend to produce higher scores on post-test. As very few data were gathered from the pilot experiment, detailed analysis could not be performed and thus little evidence to show that the teaching model is effective. However, there was indication that the teaching had some effects on the students' learning as measured by the difference in pre- and post-tests, even though it was small.

In light of these conclusions, the following procedures are proposed for the main experiment with some changes made towards the design and analysis of the main experiment:

- Student participation will remain voluntary with strong encouragement from the course lecturer.
- The use of remedial classes outside the normal lecture hours shall remain. The provision of remedial help in understanding statistical concepts and computations to students outside the normal lecture hours remains the stated aim. A small classroom which could hold up to twenty people was chosen as the venue for the experiment to allow for interactions and discussions.
- The teaching materials were developed to fit a certain amount of teaching time and in conjunction with the teaching of several concepts in eight differerent topics.
- The pre- and post-test are made comparable (but not identical) in terms of their structure and difficulty level.
- To investigate which factor, if any, was associated with the increase in learning performance of the students, a 2^3 randomised factorial block design with confounding is used.

5.5 The Main Experiments

5.5.1 Introduction

The main purpose of the experiment was to test the efficacy of the components in the teaching models and to test the effectiveness of the models in the teaching of different statistical concepts. The work involved looking at the effects of the teaching models on students' learning and that students would experience some form of learning increase following the employment of the teaching model .

The experiments were conducted in two phases. Phase 1 involved experimenting on 34 students in semester 2 of the 1994/95 academic session. These were conducted in eight consecutive weeks covering eight different topics of statistical inference. Phase 2 involved experimenting on 24 students in semester 2 of the 1995/96 session covering only the first four topics of the Phase 1 experiment. The decision to conduct the Phase 2 experiment came following the analysis of the Phase 1 experiment which revealed a contradictory interpretation of the results. This is described later in the analysis of Chapter 6. Only one person (i.e., the author) was involved in conducting the experiments. The methods are described further in this chapter and throughout the analysis of Chapter 6.

5.5.2 Framework of the Teaching Model

The basic framework of the teaching model is described. The model was divided into three main components which makes up the three factors in the design of the experiment. This is shown in Table 5.2.

Table 5.2

Basic Components of the Teaching Models
• **Teaching Concept Separate from the Worked Examples** • **Teaching Concept in the Worked Examples**
• **No Provision of Supplementary Notes** • **Provision of Supplementary Notes**
• **Classroom Assignment (individual)** • **Classroom Assignment (pairing)**

The first component of the teaching model, i.e., 'Teaching Concept' involved the teaching of several statistical concepts associated with each of the following topics:

(1) Sampling Distribution of the Means - Week 1
(2) Confidence Interval 1(large samples) - Week 2
(3) Confidence Interval 2 (small samples) - Week 3
(4) Hypothesis Testing 1 (large samples) - Week 4
(5) Hypothesis Testing 2 (small samples) - Week 5
(6) Chi-square Goodness-of-Fit - Week 6
(7) Chi-square Independence Test - Week 7
(8) Linear Regression - Week 8

These topics were chosen for several reasons: (1) It came at the right time of the syllabus after the pilot study, (2) These were among the concepts which are difficult to teach and understand (Garfield and Ahlgren, 1988; see also Table 7.3 (a)-(b) in Chapter 7).

In the first component of the teaching model, concepts were either discussed and incorporated into the worked examples or discussed separately from the worked examples. In the second component, supplementary notes were either given or not given to the students during the course of the instruction. In the third component, students either work individually or in pairs (with one group of three if necessary) in the class assignment. The aims were to investigate: (1) the effect of teaching concepts in worked example as opposed to separately from worked example, (2) the extent to which the printed notes has an effect on students' learning, (3) the effect of working alone as opposed to working in pairs, (4) the effect of any combination of the teaching components on students' learning. Any combination of the three components, say, 'pc' would involved students being exposed to the teaching of "concepts in worked examples with no printed notes given and working in pairs", which constitute one treatment combination. The other treatment combination, say, 'n' would involved the teaching of "concepts separately from the worked examples with printed notes given and working individually".

5.6 Design of the Experiment

5.6.1 Introduction

There has been some discussion on experimentation of teaching methods using appropriate designs and methodological techniques suitable to the needs of the respective studies (Keeling and Lenz, 1966; Harding et al., 1981; Davies, 1996; Marcoulides, 1990). Most of these studies used the treatment versus control group design of experiment to investigate the effectiveness of different types of teaching methods or teaching aids in teaching statistics.

In the large classroom experiment (in Chapter 4) variables such as course of study and prior knowledge were not controlled due to impracticability of randomising the students. In this experiment, the latter issue still stands and by using small group teaching would enable the researcher to exercise more control on the subjects. Moreover, small classroom teaching tends to provide a better atmosphere for learning (Borresen, 1990; Keeler and Steinhorst; 1995) where interactions among students and teacher are possible to allow for even greater learning to take place (Webb, 1982; Dees, 1991).

The approach described in this and the next chapter showed that there are other avenues which have not been tried in experimenting methods of teaching statistics particularly at the introductory level. The main issue in this chapter is the design of the experiment relative to the development and testing of the teaching model. Related issues include randomisation and control, the instructional process and the analysis of the experiment.

5.6.2 Randomisation and Control

Two important aspects of a statistical design of experiments are randomisation and control. Randomisation is the random allocation of experimental units among treatments. In the context of this study, the subjects could not possibly be chosen randomly from the population of students taking this course due to the nature of the experiment and also because of the difficulties in getting students to participate in the study. Previous results from large classroom experiment also implied that extraneous variables which may affect the results of the study need to be controlled.

The importance of randomisation was first provided by R. A. Fisher (1920). Cochran and Cox (1957) realised that equivalent groups for experimental use can be obtained by randomly assigning subjects to groups. Many variables such as sex, course of study, educational background, ability, etc may influence the outcome of the results. The purpose of randomisation is to allow us to ignore the effect of variables which are not of interest to the study, but nonetheless may influence the dependent variable and the results of the study. So the use of random allocation should average out the effects of these extraneous variables by dividing them evenly between the groups. This is to ensure that valid conclusions can be drawn from the analysis.

However, it is difficult to randomise and match subjects in all important extraneous variables especially in educational experiments. This is true for this study as well because it was not possible to choose the subjects at random. The experimental class was dubbed 'the remedial class'. It was made clear to the students at the beginning that this is not a compulsory class and the purpose was to provide extra help with their studies; students go to the experimental classes (or remedial classes) on a voluntary basis because they feel that they needed the extra help and perhaps they are motivated and enthusiastic in learning more about statistics.

Clearly, the number of students who went to the experimental classes could not be controlled. This leads to more variation in the results. After a few attempts to get students to participate in the study, it was obvious that something should be done in order to get them to come to the classes. Every effort was made in producing 'good' teaching methods and 'good' teaching materials. From the author's perspective, the term 'good' means 'effective' in a way that once they have been exposed to the teaching methods and materials, they would be interested to come to the class again.

5.6.3 Statistical Design

In the Phase 1 experiment, students from the 1994/95 academic session attended the experimental classes on Wednesday and Thursday for eight consecutive weeks. The experiment was replicated with different treatment effects being confounded in each replication. In this experiment, there were two replications. The first 4 weeks correspond to Replicate I and the balance to Replicate II. In Phase 2, students from the 1995/96 academic session attended the experimental classes on Wednesday and Thursday for four consecutive weeks only, repeating Replicate I from Phase 1. In the Phase 1 and Phase 2 experiments, a 2^3 randomised factorial block design was used. The subjects were exposed to a different treatment combination each week.

In Phase 1, eight treatment units were assigned randomly between the two days (Wednesday and Thursday) in each week and between the four weeks (Weeks 1, 2, 3 and 4) in Replicate 1. Similarly in Phase 2, the treatment units were assigned randomly between the two days in each week and between the four weeks (Weeks 1, 2, 3 and 4). For Replicate II, the treatments were assigned randomly between the days and the rest of the weeks (Weeks 5, 6, 7 and 8). The Replicate II design is applicable to the Phase 1 experiment only. The procedure and the reasons for confounding the factors which leads to the design of Replicates I and II are described.

Since the teaching model emphasised the three teaching components which are Notes (N), Pairing (P) and Concepts (C), it was felt that it is most important to estimate all the three main effects, i.e., N, P and C. Therefore, the main effects must not be confounded with weeks. In order to avoid this, based on the odds and even rule in the experimental design, factors NP, NC and PC have to be confounded with weeks. This then lead to the actual design of Replicate I in Table 5.3(a).

Table 5.3(a): Experiment Design of Replicate I

Phase 1

Weeks / Treatment/Days	Week 1 (Sampling)	Week 2 (CI 1)	Week 3 (CI 2)	Week 4 (HT 1)
1/Wednesday	nc	c	n	(1)
2/Thursday	p	np	pc	npc

*** Phase 2**

Weeks / Treatment/Days	Week 1 (Sampling)	Week 2 (CI 1)	Week 3 (CI 2)	Week 4 (HT 1)
Wednesday	pc	(1)	np	nc
Thursday	n	npc	c	p

Key: n = notes, p = pairing, c = concept

* Different randomisation of treatment factors

In Replicate I, the situation can be explained as follows.

Based on the odds and even rule, NP is estimated as

$$(n-1)(p-1)(c+1)$$

$$\Rightarrow \quad [(np + c) + (npc + 1)] - [(nc + p) + (n + pc)]$$

where $(np + c)$, $(npc + 1)$, $(nc + p)$ and $(n + pc)$ correspond to Week 2, Week 4, Week 1 and Week 3, respectively.

So, the NP estimate may also be written as

[Week 2 + Week 4] -[Week 1 + Week 3]

where NP is confounded with weeks.

Similarly PC is estimated as

$$[(n + pc) + (npc + 1)] - [(nc + p) + (c + np)]$$

or [Week 3 + Week 4] - [Week 1 + Week 2]

and that NC is also confounded since

$$NC - NP \times PC = NP^2C = NC$$

Hence, in Replicate I, the main effects N, P, C, and NPC were estimated and NP, PC and NC were confounded with weeks. For Replicate II, it was decided to adopt a different confounding scheme to obtain information on the two factor interactions. This involved confounding a main effect and the three factor interaction as well as a two factor interaction. An alternative would have been to confound a two factor interaction and two main effects. However, this was not desirable as estimation of the main effects was important.

It was found that factors concept (C) and notes (N) were important and therefore must be estimated. To avoid N and C from being confounded and again based on the odds and even rule, factors P, NPC were chosen such that P x NPC = NP^2C which resulted in factor NC being confounded as well. This then lead to the design of Replicate II in Table 5.3(b).

Table 5.3(b): Experiment Design of Replicate II
(applicable to Phase 1 only)

	HT 2	χ^2 Goodness of Fit test	χ^2 Test of Indep.	Linear Regression
Weeks / Treatment/Days	Week 5	Week 6	Week 7	Week 8
1 / Wed	np	n	(1)	npc
2 / Thurs	pc	c	nc	p

Thus in Replicate II, main effects N, C plus two interaction factors NP and PC were estimated while factors P, NPC and NC were confounded with weeks. The reasons for confounding those factors with weeks was because the design consisted of blocks of size 2 with 8 treatment units assigned randomly in between the weeks and days. The design needs 4 blocks for one replicate with 3 degrees of freedom between the weeks. As a result, three treatment combinations must be confounded in each replicate.

5.6.4 Instructional Process

In this section, the process of using the teaching model in the experiment are described. Basically, the teaching methods which were employed in the experiment linked closely with the components of the teaching models. The instructional process must be consistent with the treatment combinations assigned to respective days and weeks and different statistical topics are subjected to different types of teaching combinations (or treatments). The experiment does not only involve the implementation of the teaching methods but also the process of assessing the teaching model via students' performance. Tests were administered prior to the teaching (pre-test) and at the end of teaching (post-test). For the pre-test, students were tested based on their knowledge of the concepts taught in the lecture which they attended a week earlier prior to the experiment. The post-test was intended to see how much they have learned as a result of the experimental teaching. After the pre-test, students in both groups were given handouts stating the important key concepts, definitions and what they would be expected to know at the end of the teaching. The following present an example where different teaching combinations were employed in the experiment.

The following describes the experiment conducted in Week 1 of Replicate I which involved the teaching of several concepts in Sampling Distribution of the Means. Table 5.4 displays the teaching combination of Treatments 1 and 2.

Table 5.4: Teaching Combinations in Week 1
(Sampling Distribution of the Means)

Treatment 1 (nc)	*Treatment 2* (p)
Concepts in Worked Examples	Concepts Separate from Worked Examples
Supplementary Notes	No Supplementary Notes
Working individually on classroom assignments	Working in pairs or in group of three on classroom assignments

The instructional process began with a brief explanation of the key concepts and what was expected of the students at the end of the teaching. In Treatment 1, concepts were gradually induced in the worked example where students were constantly reminded of the key concepts at various stages of the example. A flowchart was used in aiding the explanation of the worked example. This technique was used following recommendations that a diagram can be used to introduce or review basic statistical concepts which was otherwise proved difficult to understand (Lewis and Brewer, 1992; Hawkins et al, 1992). Following the discussion of the worked example, students were asked to work individually on the class assignments which they were required to solve a few problems. The intent was to provide students with some hands-on experience in problem solving to further enhance their understanding of the concepts and computations. The assignments were collected and marked, however, no further analyses were done on it. Students also received a set of printed notes designed to help them with the exercise.

In Treatment 2, the process involved explaining the concepts separately from the worked examples. Subjects in this group were not given any printed notes apart from the worked examples sheets [see Appendix A5, items (6)-(7)] but they were asked to work in pairs on the class assignment. The experimental materials can be seen in Appendix A5.

5.7 Threats to Internal and External Validity

An important consideration in planning an experimental study is the likelihood of possible threats to the internal and external validity (Moser and Kalton, 1979; Cohen and Manion, 1992). The question is: "How likely is it that any particular threat exists in this study?" One way of assessing this likehood is to evaluate the threats on the basis of how likely they are to have an effect and plan to control for them (Keeves, 1988). If a given threat cannot be controlled, this should be acknowledged. The likelihood and threats to internal and external validity and possible ways to control the threats in this experiment are discussed.

The first threat is testing. With designs involving pre- and post-test, it is possible that the experience gained through taking the pre-test may result in an increase in the post-test score. To avoid confounding the testing effect, the use of a before and after design with control group or with all treatment groups guards against this source of invalidity is necessary. As all treatment groups take the pre-test, the testing effect is common to both groups and therefore does not explain any difference in the amount of change between them. However, this would happen if the testing effect is related to any of the factors, such that, students who subsequently worked in pairs got more benefit from the pre-test experience than students who worked individually. The before and after design would not be of much help here since each week one group get to work in pairs while the other does not.

The second threat is instrumentation. In order to evaluate the teaching models in a before and after design, two comparable tests are needed. The same test cannot be used twice over for, if it were, the testing effect might well be considerable. But if the two tests are not strictly comparable, the comparison of pre- and post-test scores is not a fair one. So, in this experiment, considerable effort was made to design the pre- and post-test materials to be made as comparable as possible, but not identical. Another aspect of the instrumentation effect occurs in the grading of the tests. For consistency, both tests were made to have equivalent number of questions with 1 point awarded for each correct answer and 0 point otherwise.

The third threat is selection. The aim of forming the treatment groups should be to make them as comparable as possible. The essence of an experimental comparison is that the groups should be as nearly as possible the same in every respect except for the independent variable. This issue was addressed in Section 5.6 where the technique of randomisation is used to permit us to ignore the effect of extraneous variables. This is thought to be the most critical issue in the conduct of this experiment. Extraneous variables such as student's ability, intelligence, attitude and social background are the possible threats to the experiment. Intelligence is accounted for by the use of pre-test and ability which is prior to the test. Finally, the ability to learn, attitude and social background are controlled through the process of randomisation.

The fourth threat is experimental mortality. Even if a high degree of comparability in the two groups chosen for the experiment is achieved, that comparability may be weakened or lost through failure to obtain the required measurements for every member of each group. In this experiment, one or two students would arrive too late to take the pre-test and another student would leave a little earlier before the post-test. To overcome this, the comparability of the two groups is recovered by calculating the pre-test average only for those who also took the post-test and vice versa. Students intermittent attendance to the experimental classes may also result in non-representativeness of the population which the experimenter seeks to generalise her findings.

The fifth threat which is concerned with external validity is the reactive or interaction effect of testing. In before and after designs, there is the danger that the pre-test may affect the responsiveness of the experimental group to the treatment. As in this case, a pre-test measuring student knowledge about a certain concept may sensitize the subjects. As a result they may be more receptive to the information about the concept and thus may produce a better result in the post-test. However, since the test questions are not identical, these would reduce the bias.

The sixth threat is the reactive effect of experimental arrangements. It is, for example, a well-known phenomenon that subjects become involved in an experiment in which they arrangements is often alluded to as the Hawthorne effect. Its importance was clearly demonstrated in the famous Hawthorne studies (Roethlisberger and Dickson, 1939) where the group were treated differently from usual and knew they were the subjects of an experiment. The effect of these arrangements was found to be so strong as to hide any effects of changes in physical conditions. In order to reduce the Hawthorne effect, the arrangement is that students go to the experimental classes which they believed to be the remedial classes. They were basically told about the experiment but were not told when the experiment will take place, the condition in which they will be in and the procedures associated with the experiments.

5.8 Conclusion

As a result of the pilot study, the experiment had to be fitted within the course teaching frame and extra time was demanded from the students outside their timetable. This led to the formation of small group remedial classes intended for students to come voluntarily for help with their studies. The overall structure of teaching and assessment materials were further improved and refined. It was decided that one person (i.e., the author) should conduct the experiment. If two lecturers of differing abilities were used, then teaching methods and teacher effects are confounded. Using different lecturers also tend to breaks up continuity of class and is likely to lead to subjects dropping out of the study. Using one lecturer intimately interested in the model may lead to preferential teaching but this is not likely in view of estimation of effects.

Then the decision to use a 2^3 randomised factorial block design with confounding was made. This was a first attempt at experimenting with methods of teaching using this type of factorial design where the effect of the main factors and interactions in the teaching model could be investigated.

To investigate other aspects of the factorial design, one could consider the following. This would be contrasted and compared with the chosen design.

Weeks

Days	1	2	3	4
Wednesday	(1)	npc	npc	(1)
Thursday	npc	(1)	(1)	npc

Key: note (n), pairing (p), concept (c)

This design will give the information about 'npc' all together as opposed to none '(1)' of the components being used. No information can be gathered on the separate components apart from these two effects [i.e., (1) and npc].

When the 2^3 randomised factorial block design was finally chosen, it was made on the basis of limited resources available. In the chosen design, it was feasible to replicate 4 weeks of two classes per week and it was two sessions per week that the experimenter could have the most. One can argue that it would have been better to have: (a) 2 weeks of four classes per week with only one effect confounded, or even better (b) 1 week of eight classes no confounding. At least 25 to 30 subjects per week are needed in (a) while at least 50 to 60 are needed to conduct the experiment in (b). Given the circumstances, this was not feasible as the issue pointed to limited resources and students' participation which coincide with the constraint in timetabling. Thus, the 2^3 randomised factorial block design had made the best use of the limited resources available.

In the literature, some studies used treatment versus control group design to compare the effects of two or more teaching methods (Keeling and Lenz, 1966; Harding et al., 1981; Marcoulides, 1990). In these studies, students' prior achievement in mathematics or statistics was used to assist with the assignment of subjects to treatment groups and to control for the differences between the groups. For example, Marcoulides (1990) used a two-factor analysis of variance to analyse factors (teaching aids or methods and prior achievement groups) and their interactions to investigate whether the differences in test performance at the conclusion of the experiment is a result of the differences in the students or due to the types of teaching aids. In contrast, the present factorial design models the post-test score using pre-test as the covariate to take account of the differences between the students in the teaching groups as well as the concepts taught in each week. The effects of the main components in the teaching model and their interactions were measured and investigated in terms of its effectiveness in teaching statistical concepts. In most previous studies, post-test alone was used to measure students' learning at the conclusion of teaching.

McCloskey et al. (1995) consider a different approach, namely, paying the students to participate in their study; even so, they found some students failed to attend regularly. There is no doubt that the present experiment is faced with the problem of subjects' participation. However, the attendance rates of some of the subjects were encouraging in spite the fact that they came outside the normal timetable hours. As it is, the teaching model was used in a natural small classroom setting and not in an artificially created environment. Therefore, if this teaching model could be used successfully in the former setting then there is a possibility that it could be adapted in another natural classroom setting.

In the small classroom experiment. the following choices were made:

- only one person should conduct the experimental teaching in order to avoid confounding of teaching methods,
- to conduct experiment in a remedial class involving small number of subjects as large experiment was found not feasible,
- subject participation should be voluntary, thus no random allocation of subjects to treatment groups, and
- to teach twice per week in eight consecutive weeks as it was impossible for subjects to come to remedial every day due to timetabling contraints.

Chapter 6

SMALL CLASSROOM EXPERIMENT - THE ANALYSIS

6.1 Introduction

Data were collected on the students who attended the classes in the two phases of the experiment. Phase 1 ran for eight consecutive weeks in the Spring session of 1994/95. Data were obtained from 34 students who were observed on multiple occasions giving 150 observations in total. Phase 2 repeated the first four weeks in session 1995/96; data were obtained from 24 students, observed on multiple occasions giving 64 observations in total. The variables involved were of the interval and ratio types of data.

Each student was assigned a number to identify them and their number of attendances was recorded. Numbers were assigned to each of the weeks beginning from Week 1, ... until Week 8. In Phase 2, the experiment involved Weeks 1, 2, 3 and 4 only. The components of the teaching model, that is, notes, pairing and concepts had two levels for each component. For notes, '1' indicates the "absence of notes" and '2' indicates "the presence of notes". For pairing, '1' indicates "no pairing or individually" and '2' indicates "pairing". For concept, '1' indicates "concept explained separately from the worked example" and '2' indicates "concept explained in the worked example". The experiment was conducted in two separate groups, each meeting once a week, either on Wednesday or Thursday. The final variables were the pre- and post-test scores. There were five questions in each test in which one mark was awarded for each correct answer and zero, otherwise. The maximum possible score for each test was 5 points.

6.2 Methods of Analysis

The main method used was the analysis of covariance. The classical purpose of the analysis of covariance is to improve the precision of the parameter estimates by taking account of variation among the experimental units that is associated with variation in the dependent variable.

In the large classroom experiment in Chapter 4, it was discovered that there were inherent differences in students' ability which were not adequately controlled. Recall that that experiment had two lecture groups. Students' participation in the experiment could not be controlled as it was decided by the class timetabling. Likewise, it was not possible to control students' attendance to specific days in the small experiment.

It was anticipated that the post-test results might be dependent on students' ability, as measured by the pre-test score. This was discovered since variation in the students' ability causes variation in the post-test results. It was assumed that students who scored highly on the pre-test were considered good and those who scored poorly on the pre-test were considered weak. So, in an effort to control for this variation, the variable pre-test will be used as a covariate. This is where covariance analysis will be used to reduce this source of variation from experimental error and adjust the treatment means for differences attributable to the covariate. The explanatory variables included the block factor - Weeks, treatment effects - Note (N), Pairing (P), Concepts (C) and the interaction effects of the treatments; 'Note by Pairing' (NP), 'Note by Concepts' (NC), 'Pairing by Concepts' (PC) and 'Note by Pairing by Concept' (NPC).

The analysis of the results is described relative to the experiments conducted in Phase 1 and Phase 2. The analysis in Phase 1 is divided into four parts. Part 1 which is the preliminary analysis will investigate the number of times each student attended the experimental classes. This section will also be looking at the differences in the pre- and post-test scores in order to investigate students' learning performance.

In Part 2, the results of fitting three logistic regression models [Models (I), (II), (III)] as shown in Section 6.3.2 are presented. First, the validity of the assumptions of the Logistic Regression Models (I), (II) and (III) was investigated using simulation plots. These are shown in Section 6.3.3. In Model (I), the analysis will discuss the effects of fitting two explanatory variables - Week and Student. The response variable for the logistic regression Model (I) is the pre-test. This model will look at the effects of the treatment week on the pre-test score taking into account students' characteristics. The intention was to investigate whether some concepts were harder than others or whether students were better prepared for some concepts than others.

Part 3 of the analysis presents another aspect of the findings using two different logistic regression models. The models will discuss the effects of explanatory variables based on the response variable, post-test. Model (II) will discuss the effects of the explanatory variables by first fitting pre-test, treatment effects - Note (N), Pairing (P), Concept (C) and the interaction effects - Note by Pairing (NP), Note by Concept (NC), Pairing by Concept (PC), and Note by Pairing by Concept (NPC). In Model (II), the analysis will not take into account the design of the experiment with confounding. The pre-test score which was also the covariate is assumed to take into account all the differences between the students based upon the post-test score. This separates out those students who have prior knowledge from those without prior knowledge of the concepts.

Model (III) will look at the effects of the treatments taking into account the design of the experiment where the differences between weeks was included in the model, which involved some confounding of factors in the model. The pre-test which was used as a covariate will look at the effect of the treatments taking into the differences between the students and the weeks. Several approaches to data analysis include the analysis of sequential deviances, tables of estimates and standard errors of the parameters, odds ratio and the 95% confidence interval of the parameters odds-ratio. Part 4 proceeds by confronting Models (II) and (III) towards investigating the factors and other possible explanations in reference to the models.

In Phase 2, the analyses are divided in three parts. In Part 1, the percentage of students with correct answers on the pre- and post-test is briefly described. The mean scores of the pre- and post-test in the first four weeks are also shown. This is followed by checking the validity of the assumptions of Models (I), (II) and (III) using simulation plots. These are shown in Section 6.4.2. Part 2 describe the results of fitting Model (I) followed by the analysis of Models (II) and (III) which are described in Part 3.

6.3 PHASE 1: Classroom Experiment on 34 students in 1994/95 session

6.3.1 Part 1: Descriptive Analysis

Table 6.1 summarises the number of times each student attended the experimental classes and performed all tests.

Table 6.1: Number of times students went to the classes and performed all tests.

Student No.	1	2	3	4	5	6	7
Number of times	6	4	5	4	8	7	6

Student No.	8	9	10	11	12	13	14
Number of times	7	5	4	6	3	5	6

Student No.	15	16	17	18	19	20	21
Number of times	8	1	1	6	6	6	6

Student No.	22	23	24	25	26	27	28
Number of times	3	4	5	2	5	4	3

Student No.	29	30	31	32	33	34
Number of occasions	1	3	3	3	3	1

Table 6.1 shows that 17 of the students attended the remedial classes on at least 5 occasions. It was also discovered that students number 5 and 15 were there all the time and students numbered 16, 17, 29 and 34 were there only on one occasion.

The percentage of students with correct answers on the pre- and post-test and the mean scores for the pre- and post-test are shown in Tables 6.2(a) and 6.2(b), respectively.

Table 6.2(a): Students with Correct Answers on the Pre- and Post-tests

Number of correct answers, x	0	1	2	3	4	5
Percentage of students with x correct answers on the pre-test (n=150)	19.3	26.7	25.3	16.7	10	2

Percentage of students with x correct answers on the post-test (n=150)	0.0	2.7	12.0	30.7	36.7	18.0

Table 6.2(a) shows that 85.4% of the students obtained at least 3 correct answers in the post-test as compared to 28.7% on the pre-test. The result indicate that students tend to perform better as a result of the teaching. However, the main intention is to see if any teaching combination gives a bigger change than others.

Table 6.2(b): The Mean Scores of the Pre- and Post-test for the eight concepts

Weeks / Mean Score	Week 1: Sampling (n=13)	Week 2: CI 1 (n=14)	Week 3: CI 2 (n=18)	Week 4: HT 1 (n=17)	Week 5: HT 2 (n=23)	Week 6: χ^2 gof (n=25)	Week 7: χ^2 ind (n=22)	Week 8: Regress (n=18)
Pre-test	2.154	2.643	1.500	0.765	1.870	1.760	1.500	2.278
Post-test	3.462	4.357	3.222	3.294	3.435	3.680	3.045	4.167

Table 6.2(b) shows that the mean post-test scores are slightly higher in Weeks 2 and 8 as compared to the other weeks. The mean pre-test scores in Weeks 1, 2 and 8 which are slightly higher than the mean pre-test in the other weeks were found to be in the score range of between 2.2 to 2.6. The large differences between the pre- and post-tests indicate that there is some form of learning increase in Weeks 4, 6 and 8. This suggests that the number of students who were thought to have learned more are higher in those weeks.

6.3.2 Logistic Regression Models

The three logistic regression models which were used in the analysis of logistic regression are described. The analysis which involved checking the validity of the Binomial assumptions of the models are also described following the description of the following models.

<u>Model (I)</u>

$$\log\left(\frac{p}{1-p}\right) = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$$

where p is the proportion of correct answers for the pre-test and $\log\left(\frac{p}{1-p}\right)$ is the logistic transformation of the proportion p.

The terms on the right-hand side (RHS) correspond to the following effects:

μ is the overall mean;

α_i is the effect of week i (i = 1, 2, ..., 8);

β_j is the effect of student j (j =1, 2,......., 34);

ε_{ij} is the residual error term.

week and student are fixed effects.

The terms on the RHS of **Model (II)** and **Model (III)** correspond to the following effects:

μ is the overall mean;

γ_{1j} is the effect of factor note j (j = 1,2);

γ_{2k} is the effect of factor pairing k (k = 1,2); γ_{3l} is the effect of factor concept l (l = 1,2);

γ_{4jk} is the note by pairing interaction; γ_{5jl} is the note by concept interaction;

γ_{6kl} is the pairing by concept interaction; γ_{7jkl} is the note by pairing by concept interaction;

ε_{ijkl} is the residual error term.

Model (II)

$$\log\left(\frac{p}{1-p}\right) = \mu + \alpha + \gamma_{1j} + \gamma_{2k} + \gamma_{3l} + \gamma_{4jk} + \gamma_{5jl} + \gamma_{6kl} + \gamma_{7jkl} + \varepsilon_{ijkl}$$

where p is the proportion of correct answers for the post-test;

α is the mean pre-test.

All treatment factors are fixed effects with pre-test as the covariate.

Model (III)

$$\log\left(\frac{p}{1-p}\right) = \mu + \alpha + \beta_i + \gamma_{1j} + \gamma_{2k} + \gamma_{3l} + \gamma_{4jk} + \gamma_{5jl} + \gamma_{6kl} + \gamma_{7jkl} + \varepsilon_{ijkl}$$

where p is the proportion of correct answers for the post-test;

α is the mean pre-test;

β_i is the effect of week i (i = 1, 2, ..., 8).

week and all treatment factors are fixed effects with pre-test as the covariate.

6.3.3 Checking the Assumption of the Models using Simulation Plots

Prior to the analysis of the models, the validity of the assumptions of the Logistic Regression Models (I), (II), and (III) were checked. Envelope plots of 19 simulated samples were drawn. The plots from the original data were compared with the envelopes.

It is anticipated that a slight curve could arise as the Binomial distribution was based on only five trials. If p is small or near to 1, then the Binomial distribution is skew and skewness manifests itself by a curve in the normal plots. As an example, the following describe the simulation process of Model (I) and the production of the envelope plot using GLIM4.0 (NAG-Glim System Release 4.0, 1995).

Stage 1: *To produce a normal plot for the original data.*

(i) Fit Model (I), with pre-test as the y-variable

(ii) This will produce the fitted values (%fv) that holds the predicted number of correct answers, values between 0 and 5.

(iii) Calculate and sort the residuals [$sort sres %rs]

(iv) Calculate the normal deviate [$nd=%nd((%cu(1)-0.5) / %nu]

(v) Plot the sorted residuals against the normal deviate.

(vi) Calculate p=%fv/5. This is the probability of getting an answer correct assuming that Model (I) is correct.

Stage 2: *To produce the envelope based on 19 simulated samples.*

(i) Simulate new variable, newpretest [$calc newpre=%bid(%sr(0), nobs, p)] This will calculate a binomial random variable or simulating based on n and p where %sr(0) will generate a random number between 0 and 1.

(ii) Let newpretest be the new y-variable and refit the model.

(iii) This will produce the new fitted (%fv) based on the newpretest variable.

(iv) Calculate and sort the new residuals.

(v) Repeat (i)-(iv) nineteen times. Calculate the maximum and minimum of the sorted residuals. Then plot the sorted residuals against the normal deviates.

The envelope plots for Phase 1—Models (I), (II) and (III) are shown in Figures 6.1(a), 6.1(b) and 6.1(c), respectively.

Figure 6.1(a): Simulation (envelope) plot for Model (I) - Phase 1

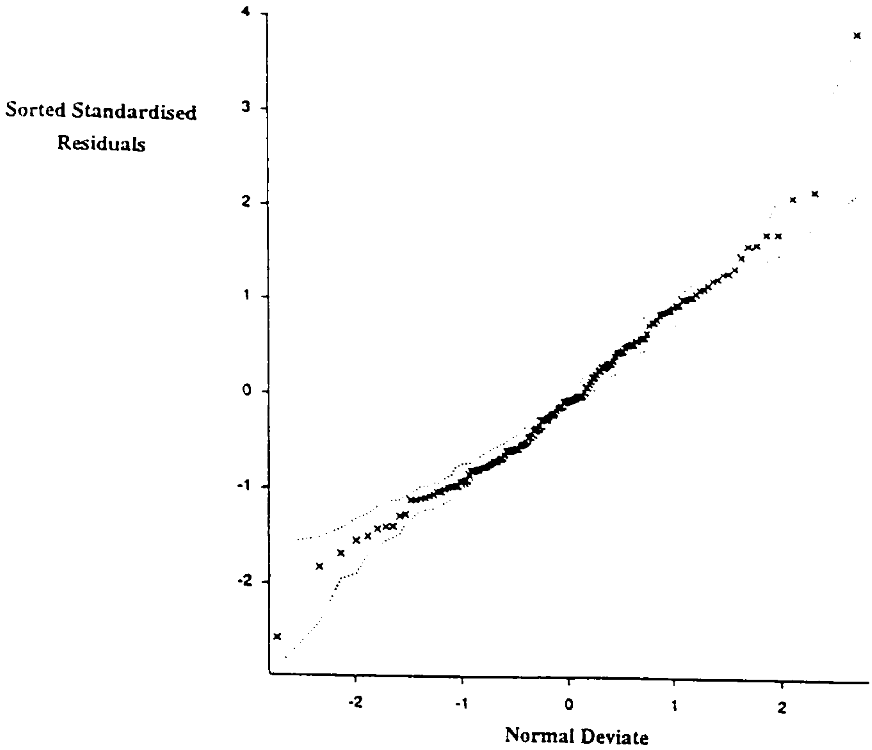


Figure 6.1(b): Simulation (envelope) plot for Model (II) - Phase 1

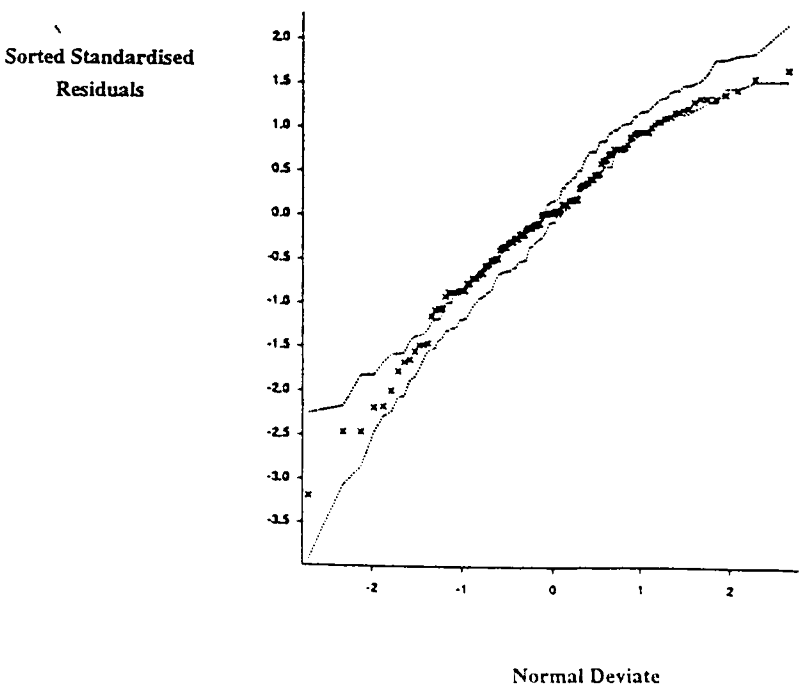
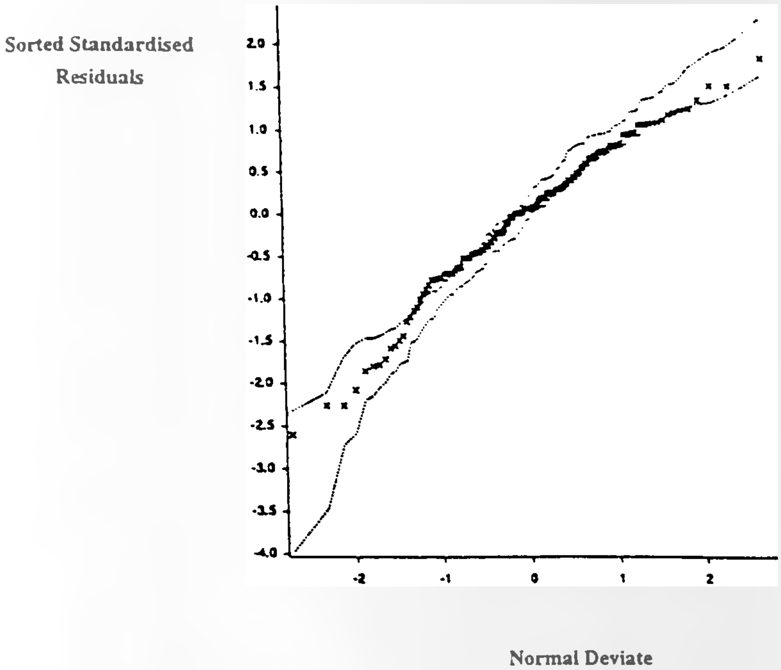


Figure 6.1(c): Simulation (envelope) plot for Model (III) - Phase 1



In Figure 6.1(a), the simulation for Model (I) of Phase 1 shows that most of the residual values lies within the envelopes. This indicates that the assumption that the model with pre-test as the dependent variable has an error distribution which is Binomial, is valid.

In Figures 6.1(b) and 6.1(c), the simulation plots for Model (II) and Model (III) show that there is a slight curve in the middle just outside the outer envelope with residuals at about -0.5. Apart from that, all the other points were within the envelope. The fact that some of the points are not within the envelope may indicate that the Binomial assumption of the model with post-test as the dependent variable is not strictly valid. However, this is not a serious lack of fit. A lot of this is a result of the discrete nature of the data which resulted in some straight horizontal lines appearing. Effectively, the assumption that the model is approximately valid enables the researcher to perform the necessary analysis on the models.

6.3.4 Part 2: Analysis of Model (I)

Several issues will be investigated. One is to investigate the differences between the weeks that would suggest some differences in the concepts. Secondly is to look at the effects of student characteristics, i.e., whether there are significant differences between the students based upon the pre-test score. Also, whether students were better prepared for some concepts relative to others. Others include checking for the level of difficulties in the concepts, i.e. whether one concept is harder than the other.

As a result of fitting Model (I), analysis of deviances (sequential and by omission from the model with all terms) and the parameter estimates for week effect are produced. The analysis of deviances and the parameter estimates for weeks can be seen in Tables 6.3 and 6.4, respectively.

Table 6.3: Analysis of Deviances for Model (I)

Source	d.f.	Reductions in Sequential Deviance	Deviances by Omission	p-value
Week	7	34.4	43.9	0.0001
Student	33	93.4	93.4	0.1311

Table 6.3 shows that the results indicate a bigger change in reduction on week from deviance by omission of full terms as compared to when it is obtained from the sequential deviances. However, the reduction remain the same for student. The p-value (0.0001) for sequential deviance indicates that there is a significant difference in the pre-test, at the 5% level, between the weeks. The difference is further investigated by looking at the parameter estimates for the weeks in Table 6.4. The estimate for Week (i) represents the difference in score between Week (i) and Week (1), which dealt with Sampling Distribution. Student effects are not significant and therefore they are not included in Model (II) and Model (III).

Table 6.4: Parameter Estimates for Week

Parameter	Topics	Estimate	Standard Error
μ	-	0.331	0.479
Week (2)	Confidence Interval 1	0.410	0.423
Week (3)	Confidence Interval 2	-0.619	0.414
Week (4)	Hypothesis Test 1	-1.791	0.461

Week (5)	Hypothesis Test 2	-0.383	0.392
Week (6)	χ^2 Goodness of Fit	-0.445	0.390
Week (7)	χ^2 IndependenceTest	-0.620	0.403
Week (8)	Linear Regression	0.254	0.427

Table 6.4 shows that Week (2) has the largest estimate of 0.410 followed by Week (8) which has an estimate of 0.254. This could indicate that concepts taught in Week (2) and Week (8) which is on Confidence Interval 1 and Linear Regression, respectively were found easy as compared to the concepts in the other weeks. On the other hand, concepts taught in Week (4) which is on Hypothesis Test 1 was found to be the hardest as indicated by the estimate.

Summary

As expected, there were significant differences among the weeks in Model (I). This indicates that there are some differences in prior knowledge associated with the different topics. It also means that pre-test should be taken into account in the analysis of Model (II) and Model (III). There was no significant effect of students' characteristics on the pre-test score. This indicates that there are no differences in prior knowledge associated with students' ability or characteristics.

6.3.5 Part 3

Section (a): Analysis of Model (II)

In Section (a), the process of analysing the data is not related to the actual design of the experiment since the logistic regression model does not take into account the week effect with no confounding of treatment factors.

The model was fitted by first adding the covariate pre-test followed by the treatment effects Note, Pairing, Concept in that order and the treatment interaction effects. The analysis of sequential deviance was produced to look at the reductions in deviance, to investigate for any significant effects of the factors when they were sequentially added to the model. Table 6.5 gives the analysis of sequential deviance for Model (II).

Table 6.5: Analysis of Sequential Deviance for Model (II)

Source	d.f.	Reductions in Deviance	p-value
Covariate-pretest	1	25.98	0.0047
Note (N)	1	1.66	0.2059
Pairing (P)	1	0.09	0.7518
Concept (C)	1	0.71	0.4028
N x P	1	2.39	0.1213
N x C	1	0.54	0.4795
P x C	1	0.10	0.7518
N x P x C	1	4.26	0.0381

When covariate pre-test was added to the model, there is a significant reduction in deviance with $p = 0.0047 < 0.05$. This indicates that there were differences in the post-test associated with the pre-test scores. For the treatment effects, the three factors interactions, NPC was found to be significant with $p = 0.0381 < 0.05$. This indicates that there were differences in the post-test scores associated with the teaching combination of Notes, Pairing and Concepts. To further investigate the efficacy of the main and interaction effects of the factors in the model, Table 6.6 displaying the parameter estimates, standard errors, confidence intervals and the odds-ratios are produced for Model (II).

Table 6.6: Parameter Estimates, Confidence Intervals and Odds-Ratio for Model (II)

Parameter	Estimate	Standard Error	Odds-Ratio	* 95% Confidence Interval $\exp(\hat{\mu} \pm 1.96\hat{s})$
μ	0.3474	0.2361	-	(0.891, 2.248)
Pretest	0.3175	0.0689	1.37	(1.200, 1.573)
Note (N)	-0.06987	0.2924	0.93	(0.526, 1.654)
Pair (P)	0.05586	0.3425	1.06	(0.541, 2.069)
Concept (C)	0.6817	0.3487	1.98	(0.993, 3.916)
N x P	-0.1055	0.4563	0.90	(0.905, 2.201)
N x C	-0.9084	0.4654	0.40	(0.162, 1.004)
P x C	-0.8622	0.5028	0.42	(0.158, 1.131)
N x P x C	1.382	0.6716	3.99	(1.068, 14.850)

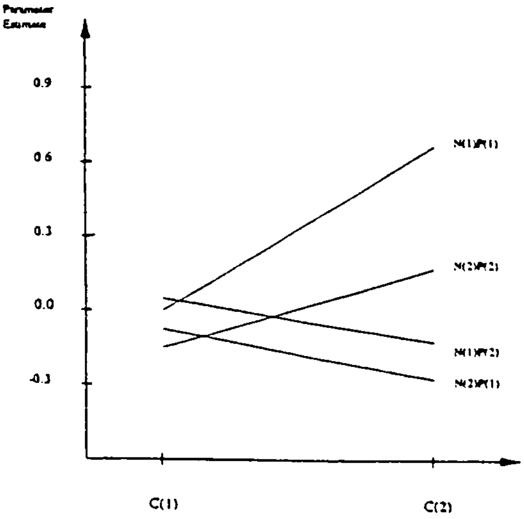
* Scale of odds

From Table 6.6, it can be seen that the interaction factor NPC is significant with a t-value equal 2.06 (> 2 s.e. from the mean). In this particular model, the pre-test score was assumed to take into account all the differences between the students and the weeks as indicated in Model (I). There was no block effect of week, hence there was no confounding of treatment factors.

In this analysis, the results indicate that factor 'C' which involved explaining the concepts in worked example has a positive influence on students' understanding when there are no notes and no pairing of students. It was found that the odds in getting correct answers doubled for those students who were taught the concepts through worked examples with a 95% confidence interval for the odds of answering a question correctly of between 0.993 and 3.916. Similarly for the interaction factor NPC, there is some evidence that the combination of the three teaching factors has some influence on students' understanding. The effect of NPC is to increase the log-odds of a correct answer on the post-test by 1.382. This means that the odds increases by a factor of 3.99 with a confidence interval of between 1.068 and 14.85. However, the interval is large which indicates that the estimates are not very precise.

The effect of NPC interaction is investigated in Figure 6.2. The response to three factors interaction of NPC shows that there is interaction between notes, pairing and concepts (NPC) on the post-test adjusted for the pre-test score. The graph suggests that there is an increase in students' learning when the method of teaching concepts in the worked examples was used, both in the presence and absence of notes and students working in pairs with N(1)P(1)C(2) providing a slightly higher parameter estimate than N(2)P(2)C(2).

Figure 6.2: Response to three factors interaction of NPC (Phase 1)



The analysis of Models (II) and (III) show that the pre-test score is significant. Table 6.7 indicates the differences in the mean post-test scores at different levels of pre-test. The mean post-test increases at pre-test levels 0, 1, 2, 3 and 4 but decreases slightly at pre-test level 5. However, this still suggests that there is a relationship between the pre- and post-test, that is, the increase in the post-test is associated with the increase in the pre-test scores, i.e. $\beta_1 > 0$ [i.e. $\beta_1 \approx 0.318$ in Model (II)].

Table 6.7: The Means and Standard Errors for Post-test at Different Levels of Pre-test

Pre-test, x	0	1	2	3	4	5
No. of students at pre-test level 'x'	29	40	38	25	15	3
Mean for Post-test	2.89	3.40	3.68	3.80	4.47	4.00
Std. Error for Post-test	0.673	0.955	1.097	1.000	0.516	1.000

Section (b): Analysis of Model (III)

In Section (b), the process of analysing the data is related to the actual design of the experiment as described in Chapter 5. The treatments consisted of the combinations of three teaching factors, N, P, and C with treatment factors being confounded and partially confounded with weeks. .

Taking into account the design of experiment, the explanatory variable week was included in Model (III) with factors NP, NC and PC being confounded with Weeks 1, 2, 3 and 4; P, NPC and NC being confounded with Weeks 5, 6, 7 and 8. It was anticipated that there would be some differences between the weeks as different concepts were taught.

Model (III) was fitted by first adding the covariate pre-test, week effect followed treatment effects——Note, Pairing, Concept in that order and the treatment interaction effects. The analysis proceed by presenting the analysis of deviance as shown in Table 6.8.

Table 6.8: Analysis of Sequential Deviance for Model (III)

Source	d.f.	Reductions in Deviance	p-value
Covariate-pretest	1	16.61	0.0004
Week	7	26.56	0.0003
Note (N)	1	1.58	0.2059
Pairing (P)	1	4.47	0.0039
Concept (C)	1	0.56	0.4386
N x P	1	0.21	0.6547
N x C	*	nil	nil
P x C	1	0.01	1.0000
N x P x C	1	0.03	1.0000

* Factor not estimated

Table 6.8 shows that when the covariate pre-test followed by the variable week were added to the model, there appears to be a significant difference ($p < 0.05$) between the weeks and the pre-test. This indicates that the differences in the post-test were associated with the pre-test and the different topics. It is noticed that the reduction in deviance for factor Pairing (P) is 4.47, which is significant. Here, the deviance is sequential and is only based on adding in pairing given pre-test, week and notes in the previous model. Referring back to the design of the experiment in Chapter 5, it was discovered that the reduction in deviance for Pairing (P) which is 4.47, comes from Weeks 1 to 4 only. Any differences between NP and P in Weeks 5 to 8 were assigned to the between weeks deviance only.

To investigate the differences between the weeks and the treatment factors, Table 6.9 displaying the parameter estimates, standard errors, confidence intervals and odds-ratio for factors in Model (III) are consulted.

Table 6.9: Parameter Estimates, Confidence Intervals and Odds-Ratio for Model (III)

Parameter	Estimate	Standard Error	Odds-Ratio	* 95% Confidence Interval $\exp(\hat{\mu} \pm 1.96\hat{s})$
μ	0.7527	0.5316	-	(0.749, 6.019)
Pretest	0.2669	0.0754	1.31	(1.126, 1.514)
Week(2)	0.9091	0.6144	2.48	(0.744, 8.273)

Week(3)	-0.0949	0.5213	0.91	(0.327, 2.527)
Week(4)	0.0405	0.5235	1.04	(0.373, 2.907)
Week(5)	0.3430	0.6168	1.41	(0.421, 4.721)
Week(6)	-0.0487	0.4606	0.95	(0.386, 2.349)
Week(7)	-0.6004	0.4611	0.55	(0.222, 1.354)
Week(8)	0.9693	0.4661	2.64	(1.057, 6.573)
Note (N)	-0.3305	0.2664	0.72	(0.426, 1.212)
Pair (P)	-0.6941	0.5151	0.50	(0.182, 1.370)
Concept (C)	0.1172	0.2647	1.12	(0.669, 1.889)
N x P	0.1221	0.7239	1.13	(0.273, 4.670)
N x C	**	nil	nil	nil
P x C	-0.3477	0.7128	0.71	(0.175, 2.855)
N x P x C	0.1763	1.093	1.20	(0.140, 10.166)

* Scale of odds; ** Factor not estimated and confounded with weeks

Adjusting for the pre-test and all treatment effects, there are higher log-odds of answering a question correctly in the post-test. The estimates for Week (2) and Week (8) were found to be larger than the estimates in the other weeks. This may indicate that Confidence Interval 1 and Linear Regression are found as easy compared to the concepts taught in the other weeks. However, only Week (8) was found to be significant with a t-value of 2.08 (> 2 s.e. of mean). Looking at the treatment factors, the estimates for the treatment factor concept (C) and the interaction factors NP and NPC were positive. On the other hand, none of them were found to be significant.

When more parameters were added to Model (III), standard error for NPC (1.093) increased invariably as compared to the standard error for NPC in Model (II) which was 0.6716. Effectively, the effect of NPC is to increase the log-odds of a correct answer on the post-test by 0.176. This means that the odds increases by a factor of 1.20 with a confidence interval of between 0.140 and 10.166. However, the interval is extremely wide indicating that the estimates are not very precise.

6.3.6 Part 4: Confronting Model (II) and Model (III)

The results gathered from Model (II) and Model (III) are discussed further in this section and the issue of the effect of weeks will be highlighted. Other factors in the conduct of the experiment which are likely to have some influence on the results will be investigated.

Clearly, the results coming from the two models are conceptually different. In Model (III), the analysis was related to the design of the experiment and all treatment effects were found not significant. When variable week was included in the model, then there is confounding. Hence, there were less sensitive estimates of treatment effects like notes, pairing, concepts and the interaction factors. Model (III) indicates that, although the pre-test score has taken into account some of the differences between students prior to the experiment, there are other differences between the weeks which were not measured by the pre-test.

In Model (II), the analysis took no account of the design of the experiment as the effects of week were ignored and there was no confounding of treatment factors. It was assumed that the pre-test score would take into account the differences between the weeks and all the differences between the students. However, there appears to be another result emerging from there. The treatment estimates are more sensitive in the sense that treatment estimates are larger and the effects were found to be significant in several cases as shown in the analysis of Model (II) in Part 3 of Section (a). More sensitive estimates of the treatments are created because the model does not take into account the differences in the concepts taught in each week.

In Model (II) when 'NPC' was estimated, effectively week was not controlled. In Replicate 1, it would not make any difference if week was included in the model or excluded from the model because the same estimate would be obtained for NPC since it was not confounded with weeks. But then, when both Replicate I and II are used, then the way NPC was estimated in Replicate II was by comparing the weeks. Effectively, the subjects were being compared as well. To look further we refer to the following defining equation for the NPC interaction,

$$NPC = [(npc) + (n) + (p) + (c)] - [(np) + (nc) + (pc) + (1)]$$

and the overall design of experiment as follows,

Week

Day	1	*2	3	4	5	6	7	*8
Wed (W)	nc	c	n	(1)	np	n	1	npc
Thurs (T)	p	np	pc	npc	pc	c	nc	p

* is for the "easy" weeks with high scores

The above defining equation can also be written in terms of days and weeks as follows

NPC = [(T4 W8) + (T1 T8) + (W2 T6) + (W3 W6)] - [(T2 W5) + (W1 T7) + (T3 T5) + (W4 W7)]
* * * *

where * denote "easy" weeks; W8 (Wed in Week 8); T8 (Thurs in Week 8); W2 (Wed in Week 2); T2 (Thurs in Week 2).

So NPC has a ratio of 3:1 in favour of the "easy" weeks and confounded with weeks in Replicate II. This means that in Model (II), part of the estimated effect of NPC comes from a difference between the weeks. On further checking the situation, it was found that the ratio 3:1 also correspond to treatment effects P, NP and PC. Treatment effects NP and P were also found to be affected by "easy" weeks and also confounded with weeks. Thus when week was included in Model (III), there was no effect of "easy and hard" weeks and thus the significance associated with NPC will diminish. So the fact that NPC is significant in Model (II) but not in Model (III) must be due to some effects of weeks. For that reason, the experiment had to be repeated in Replicate I in order to ensure that NPC is estimated within weeks.

6.4 PHASE 2: Classroom Experiment on 24 students in 1995/96 session

In this section, the analysis of the Phase 2 experiment is described. The analysis is divided in three parts. In Part 1, the number of students who attended the experimental classes, the percentage of students with correct answers on the pre- and the post-test and the mean scores of the pre- and post-test are briefly described. This is followed with the analysis of the Logistic Regression Models (I), (II) and (III) as described in Section 6.3.2. Prior to the analysis, the validity of the assumptions of the models were investigated using simulation plots. These are shown in Section 6.4.2. This is followed by the analysis of Model (I) in Part 2 and both Models (II) and (III) in Part 3.

6.4.1 Part 1: Descriptive Analysis

The total number of students who went to the experimental classes in Phase 2 and performed all tests was 24. Out of those, 7 students went to the class all the time, 6 went three times, 7 went to the class twice and 4 of them went only once. The percentage of students with correct answers on the pre- and post-test are shown in Table 6.10(a). It was found that 25% of the students obtained at least 3 correct answers on the pre-test as compared to 84.4% in the post-test.

Table 6.10(a): Students with Correct Answers on the Pre- and Post-tests (Phase 2)

Number of Correct Answers, x	0	1	2	3	4	5
Percentage of students with x correct answers on the pre-test (n=64)	14.1	31.2	29.7	18.8	6.2	0
Percentage of students with x correct answers on the post-test (n=64)	0	3.1	12.5	15.6	39.1	29.7

Table 6.10(b): The Mean Scores of the Pre- and Post-tests for the four concepts

Weeks / Mean Score	Week 1: Sampling (n=14)	Week 2: CI 1 (n=16)	Week 3: CI 2 (n=18)	Week 4: HT 1 (n=16)
Pre-test	1.357	2.375	1.556	1.563
Post-test	3.643	4.125	3.556	3.875

Table 6.10(b) shows that apart from the mean post-test scores which is slightly higher in Week (2), the mean post-test scores are more or less the same in Weeks 1, 3 and 4. However, the mean pre-test score in Week (2) is also slightly higher than the mean pre-test score in the other weeks.

6.4.2 Checking the Assumption of the Models using Simulation Plots

As mentioned earlier in Section 6.4, an attempt was made to check the validity of the assumptions of Models (I), (II) and (III) using simulation plots. The procedures carried out here were similar to the one conducted in Phase 1 experiment, as described previously in Section 6.3.3. The envelope plots for Phase 2—Models (I), (II) and (III) are shown in Figures 6.3(a), 6.3(b) and 6.3(c), respectively.

Figure 6.3(a): Simulation (envelope) plot for Model (I) - Phase 2

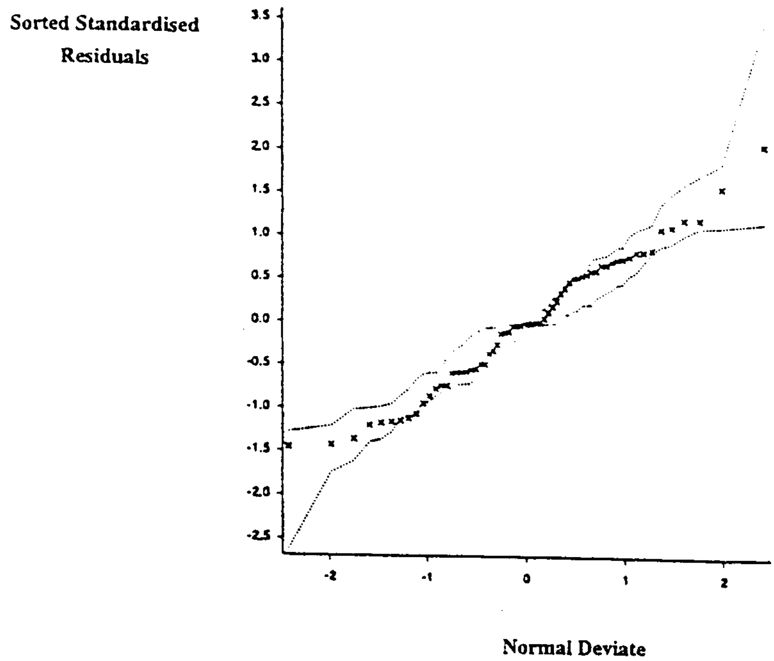


Figure 6.3(b): Simulation (envelope) plot for Model (II) - Phase 2

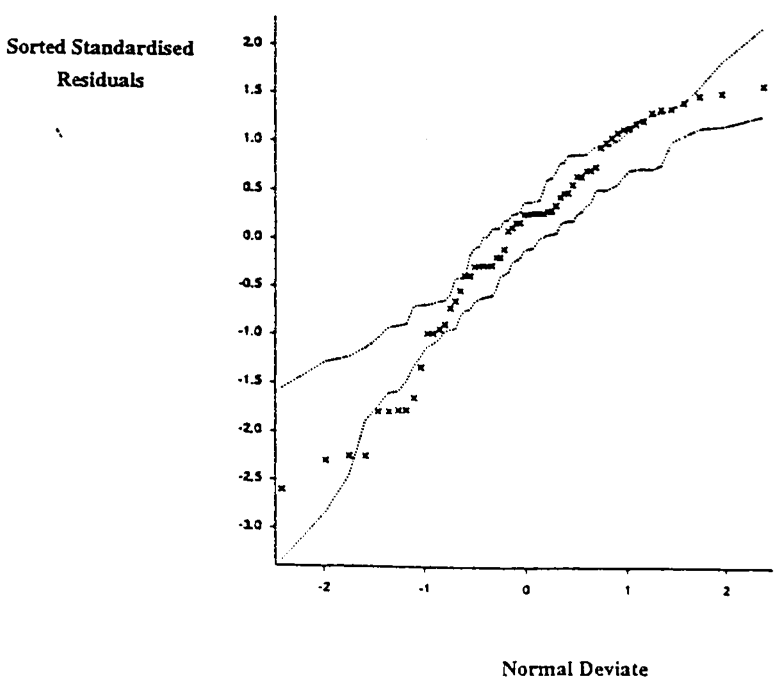
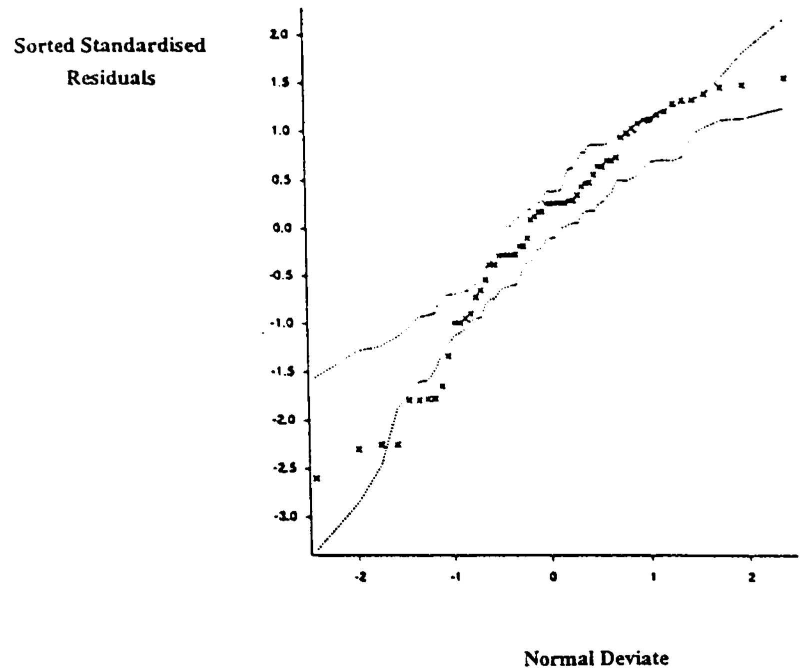


Figure 6.3(c): Simulation (envelope) plot for Model (III) - Phase 2



In checking the Binomial assumption for Models (I), (II) and (III) in Phase 2, Figures 6.3(a), (b) and (c) show that most of the residuals are within the inner envelope and that the assumptions that the models has an error distribution which is Binomial is valid. Effectively, the assumption that the model is approximately valid enables the researcher to perform the necessary analysis on the models.

6.4.3 Part 2: Analysis of Model (I)

In this section, the analysis which is similar to the Phase 1 analysis of Model (I) is described. To investigate the differences between the weeks, Model (I) as described in Section 6.3.2 was fitted. As a result of fitting the model, analysis of deviances and the parameter estimates for weeks are produced. These are shown in Tables 6.11 and 6.12, respectively.

Table 6.11: Analysis of Deviances for Model (I)

Source	d.f.	Reductions in Sequential Deviance	Deviances by Omission	p-value
Week	3	8.31	10.36	0.042

Student	23	27.5	27.5	0.256

Table 6.11 shows that the p-value (0.042) for sequential deviance indicate that there is a significant difference in the pre-test, at the 5% level, between the weeks. The results also indicate a bigger change in reduction on week from deviance by omission of full terms than when it is obtained from the sequential deviance. This indicates that there are some differences in prior knowledge associated with the different topics. The latter is also consistent with the result in Model (I) of Phase 1. It also means that pre-test will be taken into account in the analysis of Models (II) and (III) in Part 3 of Phase 2. The differences in weeks are further investigated in Table 6.12.

Table 6.12: Parameter Estimates for Week

Parameter	Topics	Estimate	Standard Error
μ	-	-1.032	0.550
Week (2)	Confidence Interval 1	1.111	0.399
Week (3)	Hypothesis Test 1	0.173	0.400
Week (4)	Hypothesis Test 2	0.278	0.424

Table 6.12 shows that Week (2) has the largest estimate of 1.111 as compared to the estimates in the other weeks. This indicates that in Week (2), every student tends to score higher both in the pre- and post-test and that Confidence Interval 1 was found an easy concept to understand.

6.4.4 Part 3

Section (a): Analysis of Model (II)

To summarise, the process involved fitting a logistic regression model when the factor weeks was excluded from the model. Analysis of deviances and parameter estimates are produced to investigate the effect of the covariate (pre-test) and the factors in the model. This is accomplished by displaying the analysis of deviances and the parameter estimates of Model (II) factors in Tables 6.13 and 6.14, respectively.

Table 6.13: Analysis of Sequential Deviance for Model (II)

Source	d.f.	Reductions in Deviance	p-value
Covariate-pretest	1	0.155	0.6547
Note (N)	1	0.274	0.5839
Pairing (P)	1	0.919	0.3428
Concept (C)	1	0.490	0.4795
N x P	1	0.050	0.7518
N x C	1	2.693	0.1003
P x C	1	1.102	0.2943
N x P x C	1	4.754	0.0285

Table 6.13 shows that when pre-test was added to the model, there is no significant reduction in deviance for the covariate. All other treatment effects are not significant apart from the NPC interaction with p-value equal 0.0285.

Table 6.14: Parameter Estimates, Confidence Intervals and Odds-Ratio for Model (II)

Parameter	Estimate	Standard Error	Odds-Ratio	* 95% CI $\exp(\hat{\mu} \pm 1.96\hat{s})$
μ	1.132	0.4544	-	(1.273, 7.561)
Pretest	-0.0948	0.1337	0.91	(0.700, 1.182)
Note (N)	-0.2541	0.5143	0.78	(0.2831, 2.125)
Pair (P)	0.3486	0.4937	1.42	(0.5385, 3.729)
Concept (C)	-0.0549	0.4973	0.95	(0.3571, 2.509)
N x P	-0.5852	0.6900	0.56	(0.1440, 2.154)
N x C	-0.1742	0.7441	0.84	(0.1954, 3.612)
P x C	-0.4553	0.7319	0.63	(0.1511, 0.979)
N x P x C	2.470	1.147	11.8	(1.248, 111.9)

* Scale of odds

In Table 6.14, the NPC interaction is found to be significant with a t-value of 2.15 (> 2 s.e.). There are no significant effects on the other factors in the model. The results indicate the three teaching combination NPC seem to have some influence on students' understanding of the concepts. The effect of NPC is also to increase the log-odds of a correct answer on the post-test by 2.470. This means that the odds increases by a factor of 11.8 with a very large 95% confidence interval of the parameter. Again, this indicate that the estimates are not very precise due to the small number of experimental subjects being measured each week. This situation may be remedied if there were more subjects in the study. Also, no relationship between the pre- and post-test was detected.

Section (b): Analysis of Model (III)

This involves fitting the model when the factor weeks was included in the model. In Phase 2, the experiment was repeated for the first replication where factors N, P, C and NPC are estimated.

In the analysis of Model (III) in Phase 1, it was found that the NPC interaction was not significant because it was confounded with weeks in Replicate 2. There is a possibility that the significance of NPC in Model (II) was due to the effect of some weeks being easier than others. So to ensure that NPC is estimated within weeks in Phase 2, Model (III) was fitted and the results are shown in Tables 6.15 and 6.16.

Table 6.15: Analysis of Sequential Deviance for Model (III)

Source	d.f.	Reductions in Deviance	p-value
Covariate-pretest	1	0.032	1.00
Week	3	3.59	0.31
Note (N)	1	0.15	0.75
Pairing (P)	1	1.17	0.27
Concept (C)	1	0.74	0.40
N x P	1	*	nil
N x C	1	*	nil
P x C	1	*	nil

N x P x C	1	4.75	0.03

* Factors not estimated and confounded with weeks.

Table 6.15 indicates that the NPC interaction is significant with a p-value of 0.03. On the other hand, there was no significant difference between weeks and the covariate pre-test. This indicates that the differences in post-test were not associated with the pre-test and the different topics. Table 6.16 display the results of Model (III) analysis.

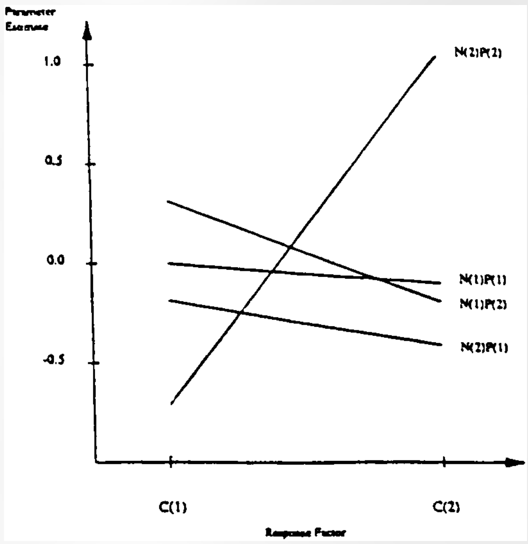
Table 6.16: Parameter Estimates, Confidence Intervals and Odds-Ratio for Model (III)

Parameter	Estimate	Standard Error	Odds-Ratio	* 95% CI $\exp(\hat{\mu} \pm 1.96\hat{s})$
μ	1.712	0.5517	-	(1.879, 16.30)
Pretest	-0.095	0.1337	0.91	(0.700, 1.182)
Week (2)	-0.3797	0.6184	0.68	(0.203, 2.301)
Week (3)	-0.0649	0.3598	0.94	(0.463, 1.897)
Week (4)	0.1406	0.3878	1.15	(0.538, 2.461)
Note (N)	-0.6338	0.3678	0.53	(0.258, 1.091)
Pair (P)	-0.1717	0.3608	0.84	(0.415, 1.708)
Concept (C)	-0.3696	0.3886	0.69	(0.323, 1.480)
N x P	**	nil	nil	nil
N x C	**	nil	nil	nil
P x C	**	nil	nil	nil
N x P x C	2.47	1.147	11.8	(1.248, 111.9)

* Scale of odds
** Factors not estimated and confounded with weeks

From Table 6.16, the estimate and standard error for the interaction factor NPC is the largest as compared to the estimates and standard error in the other factors. Factor NPC is also found to be significant with a t-value of 2.15 (> 2 s.e.). The effect of NPC is investigated in Figure 6.4. The response to three factors interaction of NPC shows that there is interaction between notes, pairing and concepts (NPC) on the post-test adjusted for the pre-test score. The graph indicated that there is some increase in students' learning when the method of teaching concepts in the worked examples was used in conjunction with notes and students working in pairs.

Figure 6.4: Response to three factors interaction of NPC (Phase 2)



The analysis of Models (II) and (III) show that the covariate pre-test is not significant. This is further investigated in Table 6.17.

Table 6.17: The Means and Standard Errors for Post-test at Different Levels of Pre-test

Pre-test, x	0	1	2	3	4
No. of students at pre-test level 'x'	9	20	19	12	4
Mean for Post-test	4.00	3.55	3.89	3.83	4.00
Std. Error for Post-test	0.866	1.146	0.994	1.115	2.000

Table 6.17 shows that the mean post-test scores are more or less comparable at different levels of pre-test. This suggests that there is no relationship between the

pre- and post-test, that is, the increase in the post-test is not associated with the increase in the pre-test scores.

6.5 Discussion and Conclusion

The aspects of analysing the teaching models using a 2^3 factorial block design with confounding have been described. Issues to be discussed are: (1) the effect of factorial design on the results, (2) the implications of study methods, and (3) the implications of experimental results.

In the experiment, students were assumed to have prior knowledge of the concepts based on the lecture they attended a week prior to the remedial classes. To control for the initial differences between the students in the groups, pre-test was used as a covariate to take account of this in the modelling process. One could argue about the effect of build up learning experiences of the students. It is recognised that the effect may exist; if there is a practice effect from the more frequent attenders, this will likely influence the before score and so, by adjusting for it in the analysis, we have implicitly taken it into account.

Vis-a- vis the choice of one person (i.e., the author) conducting the experiment, the implication is that there can be bias in teaching but this was better than having two or more persons teaching as differences between teachers would then be confounded with teaching methods. Reasonably adequate precautions were taken to reduce the Hawthorne effect. Students believed that they came voluntarily to the classes for help with their studies and not as subjects of experimental teaching. The differences between teaching methods used in the remedial classes was not likely to be detected by the subjects because different topics were taught in each week after students have been exposed to the main lectures one week prior to the remedial classes.

Results from the experiments have shown that there is still no one effective method that can be used to teach different types of statistical concepts. The findings in Phase 1 and 2 indicated that there were some differences in students' prior knowledge associated with different topics and pre-test was used to take account of the differences. An association between posttest and pretest scores was observed in Phase 1 but this was not the case in Phase 2. In the latter, there are nine cases of pretest score of 0 with a mean posttest score of 4. A possible explanation is that these students have gained little benefit from the lecture but did benefit slightly more from the remedial class.

The main findings indicated the existence of interaction between notes, pairing and concepts on the post-test adjusted for pre-test. There is some evidence that the NPC interaction is important, that when the method of teaching the concepts in the worked example is used in conjunction with the supplementary notes and students working in pairs, there is some increase in students' learning. The implication is that the combination of all three teaching components is required in helping students to learn and that a combination of two teaching components is less satisfactory. However, the use of NPC teaching model should not be viewed as the sole combination for effective teaching of statistical concepts and for improving instruction and learners' performance. The model should be used in conjunction with the main lecture, practicals (e.g. projects, demonstration, interalia) and other teaching aids (e.g., computer based learning modules). The issue on effective teaching was supported by Keeler and Steinhorst (1995) where course materials should also be organised and refined to meet the course objectives. The use of NPC model is more feasible in small remedial classes where the activity of each teaching component can easily be implemented.

Recent studies on teaching statistics also showed that there were attempts at improving course instruction and students' learning in an introductory statistics course (Yilmaz, 1996; Roiter and Petocz, 1996). However, these studies mainly provide the recommendations for developing a course or a teaching framework which are yet to be tested and implemented. In the development of the present teaching model, the intention was not only to provide recommendations to improve instruction and students' learning but the models were also tested in a natural classroom settings to gain insights into what work and does not work as a result of the implementation of the models.

The small classroom experiment could be improved. As mentioned earlier in Chapter 5, a better design would be to conduct the experiment in 2 weeks with 4 classes per week or even better 1 week with 8 classes per week. However, this was not feasible given the constraint of timetabling. Also, in Phase 2 experiment factors N, PC and NPC could be confounded to give estimates of C, P, NC and NP. Thus, all seven effects would have been estimated. As it stands, there was no information on NC at all. It was felt to be more important to repeat Replicate 1 in Phase 2 than to have a new confounding scheme in view of the results of Phase 1 experiment.

This experiment can also be also be treated as a pilot for a more advanced study. This coincides with Pfannkuch and Browns' (1996) work where a pilot study was conducted (as part of the larger study) to investigate teaching methods that could be used to develop the understanding of probability and variation. As discovered in the present experiment, when the combinations of three teaching components was employed, there are indications of some form of learning increased as measured by students' post-test score. This combination can be tested further in a new and more controlled experimental situation such as testing the effectiveness of the three teaching combinations against a control group on several topics at different times. In the experiment, problem solving was part of the activity of working in pairs but only as a practise session for students. For future work, one could propose to evaluate students' cognitive skills on this activity through the process of statistical reasoning as suggested by Chervany et al. (1977) in their evaluation framework.

Finally, those who have participated in this study have certainly benefited from the experimental teaching. Some were keen because they are enthusiastic, motivated and hardworking. Such students have continuously shown their support and interest by coming to the classes almost every week. Others were keen because they like the idea of the small classroom teaching where they can interact and be more open about their problems. The outcome of this experiment was intended to encourage educators to explore the development of this and other teaching models which can be used in teaching statistical concepts in a normal classroom setting.

Chapter 7

THE QUESTIONNAIRE SURVEY

7.1 Introduction

To support the choice of the teaching model, an investigation into the teaching and learning of statistics is necessary. This was first accomplished by conducting a questionnaire survey on school teachers and lecturers in higher education in Scotland. This complies with the objective of investigating the feasibility of using other methodological techniques in statistical education.

Among the issues highlighted in the survey of statistics teachers in the UK (Hawkins, 1989a) were respondents' own difficulties with statistics and their opinions on who should teach statistics and probability. In an earlier survey of statistics teaching in schools, Holmes et al. (1981) had also touched on the issue of how teachers acquired their knowledge of statistics. Even though these and other important aspects were investigated in both surveys, it is still not clear the extent to which their results apply to teachers and lecturers in Scotland since most of the information was gathered from the respondents in England and Wales. This survey looked at some other aspects of teaching statistics in Scotland which were not investigated in the previous surveys. The aims of the survey were:

- To look at the effectiveness of the various teaching methods used in teaching statistical concepts at CSYS level and at the higher education introductory level,
- Do teachers and lecturers find difficulties in teaching certain concepts? If so, which ones and why?
- What factors contribute to any difficulties in teaching the concepts?
- Do pupils and students find certain concepts difficult to understand? If so, which ones and why?
- What factors contribute to any difficulties experienced by pupils and students in understanding the concepts?

7.2 The Methods

The survey involved sending out two questionnaires. One is a questionnaire which was sent in mid-October 1994 to 45 schools in Scotland including 8 independent schools. These schools were chosen because they had participated in the teaching of CSYS Mathematics III (Statistics paper) in 1994 (SEB, 1994). Since the CSYS Mathematics III (Statistics paper) syllabus is equivalent to most universities' first year Introductory Statistics courses, there is a need to investigate and compare between the teaching of statistics in school and university. The second is the questionnaire which was sent to the lecturers in higher education who were currently involved in the teaching of statistics at the introductory level. Questionnaires were sent directly in early December 1994 to the lecturers in the Departments of Mathematics and Statistics in 10 universities in Scotland.

7.2.1 Schools and Local Education Authority

The majority of the schools under study are governed by the local Education Authority in nine Regional Councils. Hence, it was necessary to establish contacts with representatives from the education authorities and to inform them about the study. The Regional Councils are Strathclyde, Central, Lothian, Borders, Fife, Tayside, Dumfries and Galloway, Highland and Grampian. A letter which explained the aims of the study and the need to investigate the teaching of CSYS statistics in school was sent to the director of each Regional Council in early October 1994. The main reason was to obtain permission to conduct the study in schools. The process of sending the letter and getting the reply from each of the local Education Authority took approximately four weeks to complete. In spite of the bureaucratic process, permission was obtained rather easily from all the Education Authorities even though two of the regions were quite meticulous in the procedures. Representatives of the Education Authorities from Strathclyde, Lothian and Fife regions insisted that the researcher write to the head teacher in each school as they would have more control on the teacher's participation in the study.

As soon as permission was obtained from the director or representatives of the Education Authority, a copy of the letter from the local Education Authority, a set of questionnaires, a prepaid envelope and a cover letter were sent out to the principal teacher of mathematics via the head teacher of the schools in the respective regions. The cover letter explained the aims of the study as mentioned in Section 7.1, the interest in conducting the study in school and the return due date for the questionnaire, which was four weeks from the date on the cover letter.

The questionnaires were sent out in stages to all the schools in the nine regions following the reply from each Education Authority. The process of sending the questionnaires and getting the replies back took approximately 3 months to complete, beginning mid- October 1994. Fourteen schools responded within four weeks after the questionnaire was sent out to them followed by another ten schools four weeks later. Out of those who responded, five schools decided not to participate in the study because they no longer offered CSYS in Mathematics III (Statistics paper). A follow-up letter was sent out to those schools who did not respond the first time at the end of the third month after the questionnaire was first sent out. Two weeks later, four schools responded to the questionnaire. Out of 33 schools that responded, 28 actually completed the questionnaire. The total number of schools which did not respond at all was 12.

For the lecturers, the researcher had no prior knowledge of the actual number of lecturers teaching the courses. Therefore, a letter was sent to either the Head of the Mathematics and Statistics department or a representative at each university stating the purpose of the study. The letter was accompanied by four sets of questionnaires with a cover letter attached to each set and prepaid envelopes requesting the questionnaires be distributed among those who were involved in teaching statistics at the introductory level. Out of 40 questionnaires that were sent out, 32 lecturers from 10 universities completed the questionnaire.

7.3 The Questionnaire

The questionnaire was developed following the aims of the survey. As there was no record of similar studies being done in Scotland, the questions had to be constructed from scratch. It is recognised that any new developed instrument is subjected to criticism and this questionnaire was no exception. Thus, a pilot test was run on the questionnaire, each was sent to 2 teachers and 3 lecturers in Scotland. Upon their suggestions, further improvement was made on the overall structure of the questionnaire.

School Teachers

The questionnaire is divided into four sections: Section A (General), Section B (Teaching Statistics), Section C (Teaching Methods) and Section D (Teaching Materials). The general section covers background information about the respondent which includes their identity, schools and qualifications. Section B requests information about teachers' difficulties in teaching the concepts and the teachers' views on pupils' difficulties in understanding the concepts. Section C is intended to investigate the teaching methods used in teaching CSYS statistics and methods which they thought could help pupils to understand basic statistical concepts. Section D is intended to investigate the materials used in teaching CSYS statistics, the materials which they thought could be used to reinforce basic statistical concepts effectively and also methods of assessment which they used to measure pupils' understanding of the concepts. Questions 12 and 13 are two open-ended questions in the final section of the questionnaire. Question 12 investigates teachers' opinions as to why they think pupils find certain concepts difficult to understand and Question 13 asks them to provide any comments and suggestions about the study. A copy of the questionnaire can be seen in Appendix A7(i).

Lecturers

The questionnaire design for the lecturers is rather similar to the one for school teachers. The questionnaire was edited slightly to suit the status of higher education teaching. In Section B, more concepts were added to the existing list of concepts whereas in Sections C and D, little change was made. However, it was decided to omit the two open-ended questions from the questionnaire following poor responses on the questions from the teachers' questionnaires. On the other hand, this matter will be pursued in an in-depth interview study in Chapter 8. A copy of the questionnaire can be seen in Appendix A7(ii).

7.4 The Data

The major descriptive analysis took place mainly in Sections B, C and D. For Questions 7, 9(a), 10(a) and 11, all the response tables were based on Likert scales (Moser and Kalton, 1979) where codes were assigned to each of the response categories on a scale of 1 to 5. This is in reference to Questions 7, 9(a), 10(a) and 11 in the teachers' questionnaire and Questions 6, 7(a) and 8(a) in the lecturers' questionnaire. The assignment of codes for the respective questions are shown below.

Questions 6, 7

Category	Very Difficult	Difficult	Neutral	Easy	Fairly Easy
Code	1	2	3	4	5

Questions 7(a), 8(a), 9, 10, 11

Category	Use it often	Use it sometimes	Will likely be using it in the future	Won't likely be using it in the future	Will not use it
Code	1	2	3	4	5

The analyses were based on 28 and 32 returned and completed questionnaires from the teachers and lecturers, respectively.

7.5 Questionnaire Analysis

The analysis, which is mainly descriptive, intended to find out the respondents' view on which concepts contributed most to the difficulties in teaching and also their perception of students' difficulties in understanding the statistical concepts. The other intention was to investigate teaching methods that the respondents thought would help students to understand the concepts and also methods which they felt could effectively help to reinforce the understanding of the concepts. Tables 7.1 and 7.2 present the teachers' and lecturers' length of teaching experience and their areas of study, respectively.

Table 7.1: Length of Experience in Teaching Introductory Statistics

Length of Experience (years)	Teachers (n=28)	Lecturers (n=32)
Less than 5	15	4
5 - less than 10	5	5
10 - less than 15	4	2
15 - less than 20	1	6
20 or more	3	15

Table 7.2: Areas of Study at College or University

Areas of Study	Teachers (n=28)	Lecturers (n=32)
Statistics (Pure, Applied)	2	13
Mathematics (Pure, Applied)	10	11

Combination of Statistics and other areas	6	2
Combination of Mathematics and other areas	6	3
Engineering (Civil, Electrical)	3	1
Sciences (Chemistry)	1	2

Table 7.1 shows that slightly more than half of the teachers have less than 5 years experience in teaching introductory statistics (15). The reason could be due to the increased number of teachers teaching statistics at the CSYS level in the last five years. On the other hand, 21 (out of 32) had at least 15 years of experience. Table 7.2 shows that 13 lecturers have a degree in statistics while only 2 teachers have a degree in statistics. The majority of the teachers do not have a degree in statistics but have done statistics as part of their degree courses.

The proportion and standard errors of the topics most frequently mentioned by the teachers as being difficult to teach and the topics most frequently mentioned by the teachers as being difficult for pupils to understand are shown in Tables 7.3(a) and 7.3(b), respectively.

Table 7.3(a): The proportions and standard errors of the 10 topics most frequently mentioned by the teachers as being difficult to teach

Topic/Concepts	*** Proportion of teachers with teaching difficulties, $\hat{p}$**	**** standard error of $\hat{p}$**
Mann-Whitney Test	0.480	0.0944
Critical Region/ Significance Level	0.440	0.0938
Conditional Probability, Independence	0.370	0.0912
Confidence Interval	0.308	0.0872
Poisson	0.260	0.0828
Design of Sample Surveys	0.250	0.0818
χ^2 Test	0.280	0.0849
Significance Testing	0.240	0.0807
Null/Alternative Hypothesis	0.240	0.0807
Random Sampling	0.214	0.0775

* based on combined responses (Very Difficult and Difficult)

** s.e. of $\hat{p} = \sqrt{\dfrac{\hat{p}(1-\hat{p})}{28}}$

Table 7.3(b): The proportions and standard errors of the topics 10 most frequently frequently mentioned by the teachers as being difficult for pupils to understand

Topic/Concepts	* Proportion of pupils with understanding difficulties, $\hat{p}$	** standard error of $\hat{p}$
Conditional Probability, Independence	0.750	0.0818
Critical Region/ Significance Level	0.700	0.0866
Normal Approximation to Binomial	0.500	0.0945
Confidence Interval	0.571	0.0935
Null/Alternative Hypothesis	0.600	0.0926
Addition Law/ Multiplication Law	0.458	0.0942
Mann-Whitney Test	0.550	0.0940
Sign Test	0.350	0.0901
Poisson	0.391	0.0922
Design of Simple Surveys	0.280	0.0849

* based on combined responses (Very Difficult and Difficult)

** s.e. of $\hat{p} = \sqrt{\dfrac{\hat{p}(1-\hat{p})}{28}}$

From the tables, compared to the difficulties that teachers faced in teaching the concepts, teachers reported that a greater proportion of pupils were thought to have difficulties in understanding the concepts of conditional probabilites, critical region and significance level, null and alternative hypothesis and confidence interval. These pupils were taking the CSYS Mathematics III (Statistics paper). They were mainly good at mathematics and were highly motivated as they chose to do the subject. The teaching of this course was usually to a small group of pupils ranging between 1 to 30 in each school.

A simple linear correlation (r) and a 95% confidence interval for the correlation coefficient (ρ) was performed to investigate the strength of the relationship between teachers' difficulties in teaching and teachers' perception of pupils' difficulties in understanding the concepts. The results can be seen in Table 7.4. Six topics were chosen because these topics were among the 10 topics reported by the teachers as being difficult to teach as well as being difficult for pupils to understand.

In finding the 95% confidence interval for correlation coefficient (ρ), the researcher felt that the normal distribution is not strictly valid due to the nature of the questionnaire data which were based on ordinal scales of 1 to 5. To investigate this, the confidence intervals was first calculated using the normal z-transformation and then using the Bootstrap method in order to simulate 1000 values for ρ . Samples of size 28 were selected with replacement from the responses obtained from 28 teachers as regards the topics mentioned by the teachers as being difficult to teach (x_i) and the topics mentioned by the teachers as being difficult for pupils to understand (y_i). The correlation coefficient between x_i and y_i are calculated and this was repeated 1000 times which produced 1000 values for ρ . The Bootstrap was used to obtain the sampling distribution for r and to check that the normal assumption was valid.

Table 7.4 shows that in both cases, the correlation coefficient (ρ) are contained within the intervals with a 95% confidence but are mostly geared towards the upper end of the interval. Even though the normal assumption was not strictly valid, however, there was not much difference between the intervals as shown by both methods. The fact that r is quite large indicates that the harder the teacher found in teaching the concepts, the harder he or she thought it was for the students to understand the concepts or vice versa.

Table 7.4: The correlation between the topics most frequently mentioned by the the teachers as being difficult to teach and the topics most frequently mentioned by the teachers as being difficult for pupils to understand (r) and the 95% confidence interval for the correlation coefficient, ρ

Concepts	* Correlation coefficient, r	(1) 95% C.I. for ρ	(2) 95% C.I. for ρ
Critical Region and Significance Level	0.779	(0.567, 0.894)	(0.479, 0.900)
Mann Whitney Test	0.771	(0.553, 0.967)	(0.559, 0.928)
Null and Alternative Hypothesis	0.727	(0.479, 0.867)	(0.463, 0.889)
Poisson Distribution	0.720	(0.468, 0.864)	(0.487, 0.866)
Significance Testing	0.661	(0.375, 0.832)	(0.472, 0.890)
Confidence Interval	0.658	(0.371, 0.830)	(0.433, 0.883)

* based on all ratings (Very Difficult, Difficult, ..., Easy)

(1) Using Z transformation; $Z = \frac{1}{2}\ln\left(\frac{1+r}{1-r}\right)$

(2) Using Bootstrap 1000 simulations for ρ values

Tables 7.5(a) show the proportions and standard errors of the 10 topics most frequently mentioned by the lecturers as being difficult to teach and Table 7.5(b) show the 10 topics most frequently mentioned by the lecturers as being difficult for students to understand. The tables show that lecturers who reported having difficulties in teaching the concepts of critical region and significance level, z-test, t-test and random events and theoretical probabilities also reported that students have difficulties in understanding the same concepts.

Table 7.5(a): The proportions and standard errors of the 10 topics most frequently mentioned by the lecturers as being difficult to teach

Topic/Concepts	* Proportion of lecturers with teaching difficulties, $\hat{p}$	** standard error of $\hat{p}$
Conditional Probability, Independence events	0.600	0.0866
Critical Region, Siginificance Level	0.482	0.0883
Sampling distributions of means	0.407	0.0868
Significance Testing: Z-test	0.407	0.0868
Significance Testing: t-test	0.380	0.0858
Poisson approximation to binomial	0.360	0.0849
Random events, theoretical probabilities	0.357	0.0847
Normal approximation to binomial	0.346	0.0841
Null and Alternative Hypothesis	0.345	0.0840
Confidence Limits for population means and proportions	0.321	0.0825

* based on combined responses (Very Difficult and Difficult)

** s.e. of $\hat{p} = \sqrt{\frac{\hat{p}(1-\hat{p})}{32}}$

Table 7.5(b): The proportions and standard errors of the 10 topics most frequently mentioned by the lecturers as being difficult for students to understand

Topic/Concepts	* Proportion of students with understanding difficulties, $\hat{p}$	** standard error of $\hat{p}$
Conditional Probability, Independence Events	1.000	0.0530
Critical Region, Significance Level	0.840	0.0648
Sampling distribution of the means	0.833	0.0659
Siginificance Testing: T-test	0.808	0.0696
Normal Approximation to Binomial distribution	0.792	0.0717
Addition Law, Multiplication Law	0.778	0.0735
Confidence Limits for population means and proportions	0.750	0.0765
Significance Testing: Z-test	0.750	0.0765
Mean and Variance of the sum and difference of independent variables	0.727	0.0788
Random events and theoretical probabilities	0.720	0.0794

* based on combined responses (Very Difficult and Difficult)

** s.e. of $\hat{p} = \sqrt{\dfrac{\hat{p}(1-\hat{p})}{32}}$

Similarly for the lecturers, a simple linear correlation coefficient (r) and a 95% confidence intervals for the correlation coefficient (ρ) was also calculated to investigate the strength of the relationship between lecturers' difficulties in teaching and lecturers' perception of students' difficulties in understanding the concepts. In this case, four topics were chosen because these topics were among the 10 topics reported by the lecturers as being difficult to teach as well as being difficult for pupils to understand.

The 95% confidence intervals for ρ using the z-transformation and Bootstrap method to simulate 1000 values for ρ are shown in Table 7.6. Samples of size 32 were selected with replacement from the responses obtained from 32 lecturers as regards the topics mentioned by the lecturers as being difficult to teach (x_i) and the topics mentioned by the lecturers as being difficult for students to understand (y_i). The results indicate some positive correlation between lecturers' difficulties in teaching and their perception of students' difficulties in understanding the concepts. The correlation coefficient (ρ) are contained within the intervals with a 95% confidence. The fact that r is quite large indicates that the harder the lecturer felt in teaching the concept, the harder he or she thought it was for the students to understand the concepts or vice versa.

Table 7.6: The correlation between the topics most frequently mentioned by the lecturers as being difficult to teach and the topics most frequently mentioned by the lecturers as being difficult for pupils to understand (r) and the 95% confidence interval for the correlation coefficient, ρ

Concepts	* Correlation coefficient, r	(1) 95% C.I. for ρ	(2) 95% C.I. for ρ
Critical Region and Significance Level	0.675	(0.420, 0.831)	(0.446, 0.841)
Random Events and Theoretical Probabilities	0.653	(0.388, 0.818)	(0.319, 0.843)
Signifiance Test: Z-Test	0.641	(0.370, 0.811)	(0.395, 0.803)
Signifiance Test: t-Test	0.540	(0.229, 0.751)	(0.257, 0.761)

* based on all ratings (Very Difficult, Difficult, ..., Easy)

(1) Z transformation; $Z = \frac{1}{2}\ln\left(\frac{1+r}{1-r}\right)$

(2) Bootstrap 1000 simulations for ρ values

Table 7.7(a) is constructed to investigate the teaching methods which were used by the respondents in their teaching,

Table 7.7(a): The percentages of teachers and lecturers who often and sometimes used these teaching methods

	Teachers (n=28)		Lecturers (n=32)	
Teaching Methods	*Use it often*	*Use it sometimes*	*Use it often*	*Use it sometimes*
Lecture	64.3	28.6	90.6	9.4
Tutorials	48.0	40.0	71.9	28.1
Problem Solving	16.0	60.0	3.5	41.4
Investigative Approach	7.4	85.2	3.3	60.0

Individual Project	4.2	62.5	16.7	60.0
Group Project	4.3	69.6	10.0	40.0
Classroom Practicals	4.5	77.3	31.0	48.3
Computer Exercise	-	73.7	65.6	34.4

Table 7.7(a) shows that the majority of the respondents used the lecture method in their teaching. Computer exercises also seem to be among the methods used often by the lecturers. Apart from that, other teaching methods such as problem solving, individual projects, group projects, and classroom practicals were sometimes used by all the respondents in their teaching of the subject.

Table 7.7(b) shows that an investigative approach, classroom practicals, lectures, tutorials and computer exercises were favoured by the majority of the teachers and lecturers to be the methods that could help students to understand statistical concepts. Even though problem solving methods were used now and then by the teachers, only 35.7% said that they thought the method could help students to understand the concepts. Group projects was agreed by 59% of the lecturers to be the method that would help students to understand concepts. On the other hand, 32% of the teachers thought otherwise.

Table 7.7(b): The percentages of teachers and lecturers who thought these teaching methods could help students to understand the concepts

Teaching Methods	Teachers (n=28)	Lecturers (n=32)
Investigative Approach	67.9	62.5
Classroom Practicals	60.7	68.8
Lecture	57.1	56.3
Tutorials	50.0	65.6
Computer Exercises	50.0	65.6
Individual Project	39.3	53.1
Group Project	32.1	59.4

Problem Solving through small group discussion	35.7	56.3
Large group discussion	7.1	34.4

Table 7.8(a) shows the materials that teachers and lecturers used in their teaching. It can be seen that at least 80% of the teachers and lecturers preferred to use their own teaching notes, past exam questions, recommended textbook and problem sheets in their teaching. It was also noticed that in general, more lecturers preferred to use case studies, classroom practical teaching materials and published journals and articles in their teaching than the teachers in their teaching.

Table 7.8(a): The percentages of teachers and lecturers who often and sometimes used these teaching materials in their teaching

	Combined % (often/sometimes)	
Teaching Materials	Teachers (n=28)	Lecturers (n=32)
1. Own Teaching Notes	96.4	100.0
2. Past Exam Questions	92.9	93.3
3. Problem Sheets	92.3	100.0
4. Materials gathered from courses and workshops	90.5	46.4
5. Recommended Text	85.7	87.1
6. Support Materials from SEB	85.0	*na
7. Teaching Materials developed for teachers/lecturers	83.3	50.0
8. Classroom practical teaching materials	67.0	61.3
9. Published journals and articles	50.0	55.2
10. Case Studies	23.1	80.0

* na (not available)

Table 7.8(b): The percentages of teachers and lecturers who thought these teaching materials could be used to reinforce statistical concepts effectively

Teaching Methods	Teachers (n=28)	Lecturers (n =32)
Recommended Textbook	57.1	62.5
Problem Sheets	50.0	81.3
Past Exam Questions	42.9	62.5
Own Teaching Notes	35.7	62.5
Case Studies	32.1	68.8
Classroom Practical Teaching Materials	35.7	71.9

Table 7.8(b) shows that more than 50% of the lecturers thought that the above teaching materials could be used to reinforce the concepts effectively. For the teachers, even though they considered their own teaching notes to be the most important teaching material, 64.3% of them do not think that they could be used to reinforce the concepts effectively. This could have been because the teachers felt that pupils will have been exposed to the notes anyway during the teaching. Also, perhaps that teachers felt that pupils would need more than just the notes to have the concepts reinforced on them. However, in most of these cases, more lecturers agreed that these teaching materials can be used successfully if time and effort are devoted to it.

The final closed-ended question with regard to teachers' assessing their pupils' understanding of the concepts, 94% of the teachers mentioned that they are not likely to evaluate their pupils' understanding by giving a test after each lesson.

7.6 Discussion and Conclusion

This survey suggested that some of the opinions on statistical concepts and the teaching methods could be considered in supporting the choice of the teaching model. This chapter has also demonstrated the feasibility of conducting a questionnaire survey to investigate the teaching and learning of statistics in Scotland.

For the teachers, the sampling frame comprised of those who were teaching statistics as the CSYS level in 1994. The sample of teachers (n=28) fairly represents the population of teachers who were teaching statistics at the CSYS level. Likewise, the lecturers who responded to the questionnaire were among those who were directly involved in teaching statistics at the introductory university level and who have expressed their interest in the study. The issue here is of the sampling frame and the response rate.

In Hawkins' (1989a) survey, the sample was drawn from the mailing-list of those teachers and lecturers from 148 schools (primary and secondary) who had expressed some interest in a teaching statistics competition. Using a linked sampling procedure, the sample was extended up to three of the colleagues of the initial contact. The response rate was in the order of 35% of the institutions contacted. On the other hand, in the present survey a predefined population was contacted; a school where CSYS Statistics was taught with the teacher responsible for teaching the course as the target. As the average number of teachers teaching CSYS Statistics per school is one, the response rate which was 70% (28 out of 40 schools) is considered high. Hawkins had a self-selected sample and suggested there was a possibility of bias. In the present survey, the sample was not self-selected apart from the non-response which is an issue in all postal questionnaire surveys.

In the present survey, 10 of the teachers (n=28) have a degree in Mathematics and 6 others studied mathematics as part of a degree requirement. The impression is that it is taken for granted that mathematics teachers will take on the teaching of CSYS statistics since it is included in the CSYS arrangement for mathematics syllabus. Hawkins' survey found that at least 50% of the upper secondary mathematics teachers (n=44) favoured the mathematicians for teaching statistics and probability. However, in the present study the researcher did not investigate whether the mathematics teachers thought they should teach statistics. If this was investigated, it would make a further contribution to knowledge.

Teachers' knowledge in mathematics is often taken as a guide to their ability in teaching statistics even though in Holmes et al. (1981) survey, there was ambiguity in terms of how mathematics teachers acquired their knowledge in statistics. Several institutions of higher education also tend to place much reliance on students' achievement in mathematics, where their knowledge is often taken as a guide to their ability in understanding statistics. In fact, there is a lot of interest in teaching statistics without placing a great reliance of mathematics. For example, Loyer (1987) illustrates such concept as variability through classroom practicals to illustrate the variations in different samples based on a classroom data. Concepts such as sampling distributions were illustrated by Ricketts and Berry (1994) using a computer-intensive approach where mathematical formulation of statistical methods had been avoided.

The proportion of teachers who frequently mentioned the topics as being difficult to teach differ from the lecturers even though the common area of difficulty was found to be statistical inference. In contrast to Hawkins' (1989a) survey which investigated the three different areas of difficulty (i.e., techiques, applications and scientific methods), the present survey investigated in detail the topics which were perceived by the respondents as being difficult to teach and for students to understand. There were some correlations between these two, however, it is still not known of the factors which led to some of the variables being highly correlated. The respondents who considered the statistical inference topics as being difficult to teach could have been associated with those who also experienced some difficulties in the three areas of Hawkins' survey. However, this association needs to be further investigated prior to any firm conclusions.

As regards the teaching methods, a majority of the respondents (both teachers and lecturers) used the lecture method. This coincides with Gibbs and Harlands' (1987) study findings that lectures remained the most dominant mode of teaching in colleges of higher education. Many recognised the practical method and the investigative approach to be among the methods which could help students to learn concepts effectively. However, for some reasons teachers were quite reluctant to use them. One of the reasons could be due to teachers lack of necessary statistical training to teach them. In the training of teachers to teach statistics (Hawkins, 1990), one of the issues for concern is that there are not enough teachers with formal training in the pedagogy of teaching statistics at the school level. Even with the current change in the assessment methods which include practical coursework at the school level could not convince some of the teachers to change their practice and improve their skills in teaching practical statistics.

Finally, an attempt was made to analyse the two open-ended questions in the teachers' questionnaire. The preliminary analysis involved categorising the responses under several headings which resulted in several factors that contribute to teachers difficulties in teaching being identified. Apart from the lack of response, the analysis could not be pursued in great detail due to the variation in the responses given. Among the limitations of the questionnaire survey as described in Moser and Kalton

(1993) is that, " there is no opportunity to probe beyond the given answer, to clarify an ambiguous one, to overcome unwillingness to answer a particular question or to appraise the validity of what a respondent said in the light of how he or she said it." As such, the interview technique which shall be pursued in the next chapter is considered as the best defence against this.

Chapter 8

THE INTERVIEW: METHODS AND ANALYSIS

8.1 Introduction

As little information could be gathered from the open-ended questions in Chapter 7, it was decided to obtain more information on the teaching and learning of statistical concepts by conducting interviews of school teachers and lecturers in higher education in Scotland. The interview was intended to gain insights on which aspects of teaching and learning of statistics are important to support the choice of the teaching model. The approach is quite similar to Pfannkuch and Brown's (1996) pilot investigation where interview was used (as part of a larger study) to give an insight on what aspects of probability and variation should be taught and how they should be taught. Likewise, Konold (1989) had also used data from individual interviews in addition to questionnaire responses to develop items that could assess students' understanding of various probabilistic and statistical concepts. The present interviews, which were conducted face to face (between the author and the respondents) using open-ended questions, [see Appendix A8(i)] took about 6 months to complete.

This chapter will review the process of the interview and data collection followed by the techniques used in the analysis before presenting the results of the three conceptual frameworks. The aims of the interview are divided into three parts.

Part 1: To investigate

- factors that contribute to teachers' and lecturers' difficulties in teaching certain concepts and why;
- factors which the subjects thought would contribute to pupils' and students' difficulties in understanding certain concepts and why;
- students' misunderstanding and misconceptions and which concepts are mostly affected by the misunderstanding.

Part 2: To investigate

- teaching methods that teachers and lecturers used in teaching concepts;
- the purpose of using the methods;
- the implications of the methods used on their pupils and students.

Part 3: To investigate

- the signpost for pupils' and students' misunderstanding of the concepts;
- the signpost for pupils' and students' understanding of the concepts;
- the methods that teachers and lecturers used in evaluating pupils and students' understanding and misunderstanding of the concepts.

The interview responses were first analysed using Qualitative Data Analysis (QDA) techniques which was later combined and used together with multidimensional scaling (MDS) and cluster analysis techniques.

8.2 Procedure and Structure of the Interview

There were 16 teachers from Strathclyde, Lothian, Central, Fife and Tayside regions who had agreed to be interviewed. These were the teachers who had responded earlier to the questionnaire and indicated their willingness to be interviewed. Interview questions were constructed in conjunction with the aims of the interview. The questions which were forwarded to the interviewees (or subjects) did not necessarily follow the sequence in which they were prepared. The interview questions that were posed to them also depended on how they had responded to the open-ended questions in the questionnaire survey.

For example, Teacher A may have been asked, *why did he or she find the concept of significance testing difficult to teach* and the next question would ask the teacher *to describe factors that he or she thought contributed to the difficulties.* On the other hand, Teacher B *may not have encountered with any difficulties in teaching the concept.* For that matter, Teacher B may have been asked, *"Why do you find the concept of significance testing an easy concept to teach and why?"* During the interview, subject's responses to the preceding question prompted the researcher to ask the next question. For example, when Teacher C mentioned that *he or she finds the investigative approach very useful but was quite reluctant to use the approach;* this prompted the researcher to ask *why he or she was reluctant to use the approach.*

The 15 lecturers who were chosen from a list of lecturers at 5 universities and 2 colleges in Scotland were asked for an interview. These lecturers were among those who were involved in the questionnaire survey. They were lecturers from the University of Strathclyde, Edinburgh University, Glasgow University, University of Paisley, Glasgow Caledonian University, Clydebank and Stow College. Consent was obtained from the lecturers by telephone and an appropriate time for the interview was arranged.

For the lecturers, there was a slight change in the interview procedure. The interview was conducted as the questionnaire was sent out to them. In most cases the interviewer had no prior knowledge of the subjects on the aspect that the lecturer found difficult to teach. Hence the manner in which some of the questions were asked had to be changed slightly and probing questions were often used. For example, Lecturer A might be asked, *"Name one concept which you find difficult to teach and why"* instead of *"Why do you find the concept of confidence limits difficult to teach?"* The latter indicates that the interviewer had prior knowledge about Lecturer A's difficulties in teaching one of the concepts whereas the former question assumes otherwise.

It is recognised that subjects were chosen from the list of teachers and lecturers who have responded to the questionnaire from the stated regions. This would presumably tend to lead to a selection bias as respondents in the other regions were ignored. This is a valid criticism in any interview study. However, this study was intended to be small scale and part of the intention was to investigate the feasibility of conducting interviews to investigate the teaching and learning of statistics.

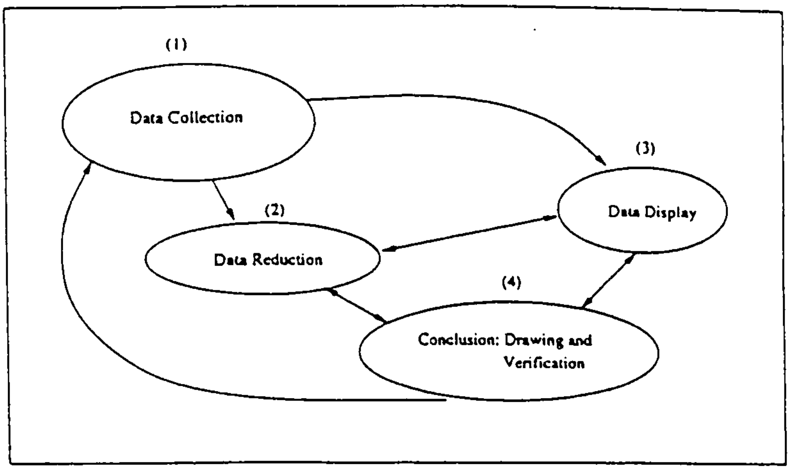
The data for this analysis comprised of the 28 interview responses which were tape recorded and transcribed. Each interview took approximately 40 minutes to complete and between two to three hours to transcribe. Three of the recorded interviews (2 teachers and 1 lecturer) could not be transcribed due to faulty recording. A sample of an interview transcript can be seen in Appendix A8(ii).

8.3 Qualitative Data Analysis (QDA)

QDA is used by social scientists among them to study human social behaviour (Harre and Secord, 1973; Huberman, 1978; Lortie, 1975), conversation and communication (Heritage, 1988) and in educational studies to study opinions, attitude and behaviour towards a certain programme, body or institution (Havelock et al. 1983; Huberman, 1981(a), Miles and Huberman, 1994). QDA was chosen for the interview study because it is suitable for processing interview transcripts and provides methods that could reveal findings which could not be elicited by quantitative methods. A brief description of QDA is given next.

In QDA, the three streams of research activities will be employed, that is, data reduction, data display and conclusion drawing (see Figure 8.1) The diagram shows that the three research activities (2), (3) and (4) and the activity of data collection (1) itself form an interactive, cyclical process. The process involved moving among the four nodes, (1) during data collection and then shuttle (2) among reduction (3) display and (4) conclusion drawing and verification.

Figure 8.1: Components of Data Analysis - Interactive Model



Source: Miles and Huberman (1994)

In QDA, the analysis will be guided by Conceptual Frameworks and Research Questions associated with each framework. For the purposes of assisting with the description of QDA, the methods using one Conceptual Framework, i.e., "Teaching and Learning Difficulties" will be illustrated. This has a number of associated Research Questions and one is

"What are the factors that the teachers and lecturers thought would contribute to pupils' and students' difficulties in understanding the concepts?"

The process of data reduction is:

1. Go through each of the 28 transcripts to look for responses that match the Research Question.
2. Different subjects may respond to the question differently. In order to differentiate between the responses, colour codes and symbols (such as #, *, etc.) were used on the transcripts. This was done mainly for personal reference. However, the following example illustrates the process.

   Teacher 3 (T3) said that, *"Pupils have not been exposed to any* *kind of statistics syllabus before"* which is related to prior knowledge in statistics. On the other hand Teacher 6 (T6) said that, *"Some are not able to think through the ideas quickly"* which is related to pupils' natural ability. In this case, the response from T3 was marked with colour code 'red' while T6 response was marked with colour code 'blue'.

3. These responses were categorised accordingly and transformed into summary phrases and quotes where they were displayed on a new sheet of paper. As a result, there were about 3 to 4 sheets of text information consisting of summary phrases and quotes which correspond to the Research Question.
4. Read through the text sheets and cross-reference with the interview transcripts to confirm certain information.
5. The text information pertinent to the Research Question will be used in the second stage of data analysis, i.e. data display.

The process of data reduction was repeated for the other research questions as well. In essence this process involves repeatedly going through the transcripts for relevant sentences or phrases.

After data reduction, the process continues to the second stage of analysis activity which is data display. Generally, a data display is an organised, compressed assembly of information that permits conclusion drawing and action. Miles and Huberman (1984; 1994) used displays to help them to understand what is happening and to do something—either analyse further or take action—based on that understanding. There are a number of different displays but the simplest is effectively just a list of the subjects together with their associated phrases extracted from the transcripts. This can be thought of as a qualitative data matrix where the rows correspond to the subjects and the columns to the research questions. The entries in the cells are the phrases extracted from the transcript which are relevant to the research question. It is at this second stage that the analysis leads to new findings. The organisation of the data into a structured framework assists with this. The analysis includes looking for similar responses to the research questions among the subjects and identifying the possible relationships between the research variables.

The third stream of analysis activity is conclusion drawing and verification. The final conclusions may not appear until the process of data collection, data reduction and data display is over. In this interview study, conclusions are drawn as the analysis proceeds usually within the second research flow of activity, the data display. The meanings emerging from the data and the conclusions are confirmed and verified with the aims of the interview, the research questions and cross-referral of the interview transcripts.

8.3.1 Conceptual Framework and Research Questions

Conceptual frameworks and research questions are the best defence against overload (Miles and Huberman, 1994). Hence, the following steps are taken to curb the overloading problems. Three conceptual frameworks were constructed as this would allow the researcher to investigate the concepts from three different perspectives:

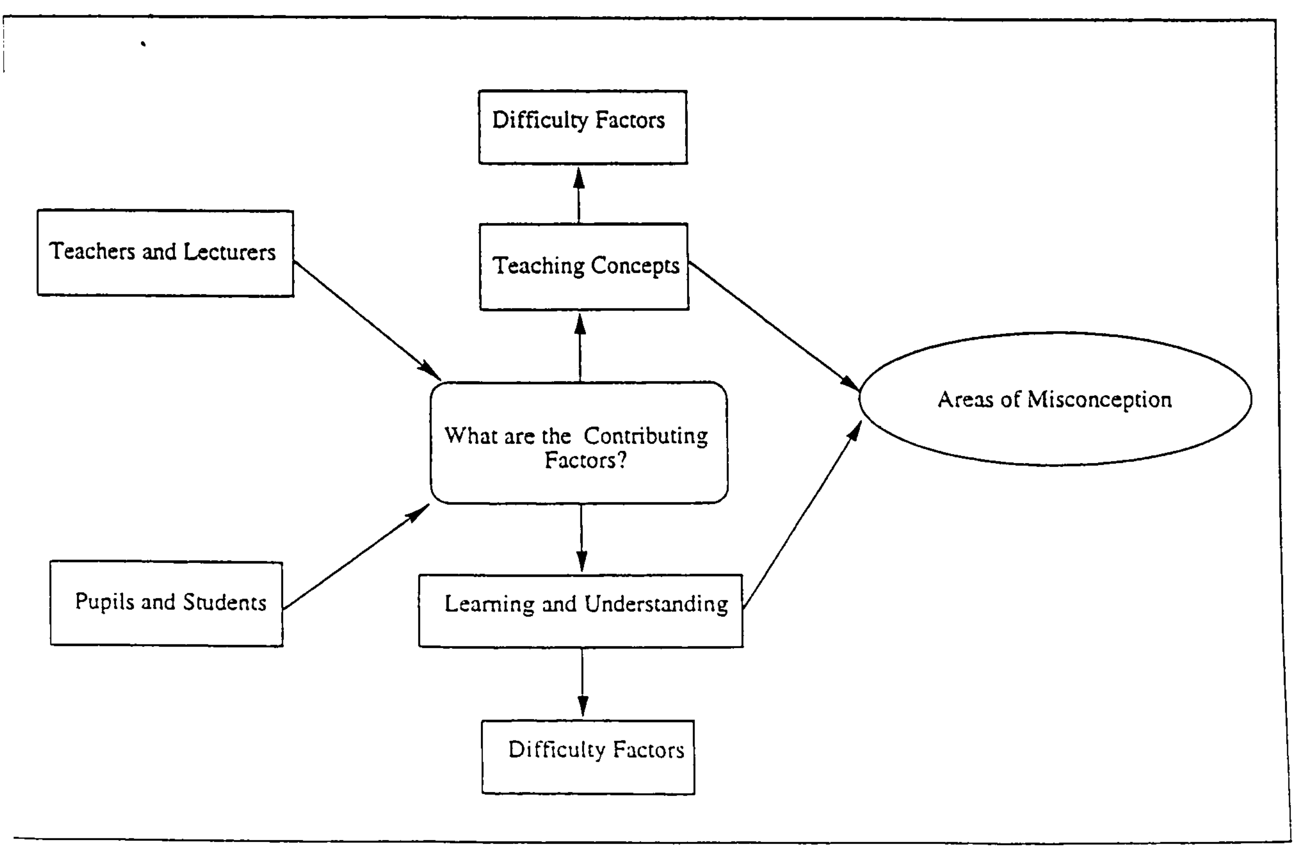
(1) Teaching and Learning Difficulties.

(2) Teaching Methods.

(3) Understanding and Misunderstanding of the Concepts.

Each of the conceptual frameworks is accompanied by several research questions. To illustrate the research activities, Figures 8.1(a), (b) and (c) together with the associated research questions were constructed. The arrows in the diagrams indicates the flow of the research activities and the direction in which the investigation was carried out.

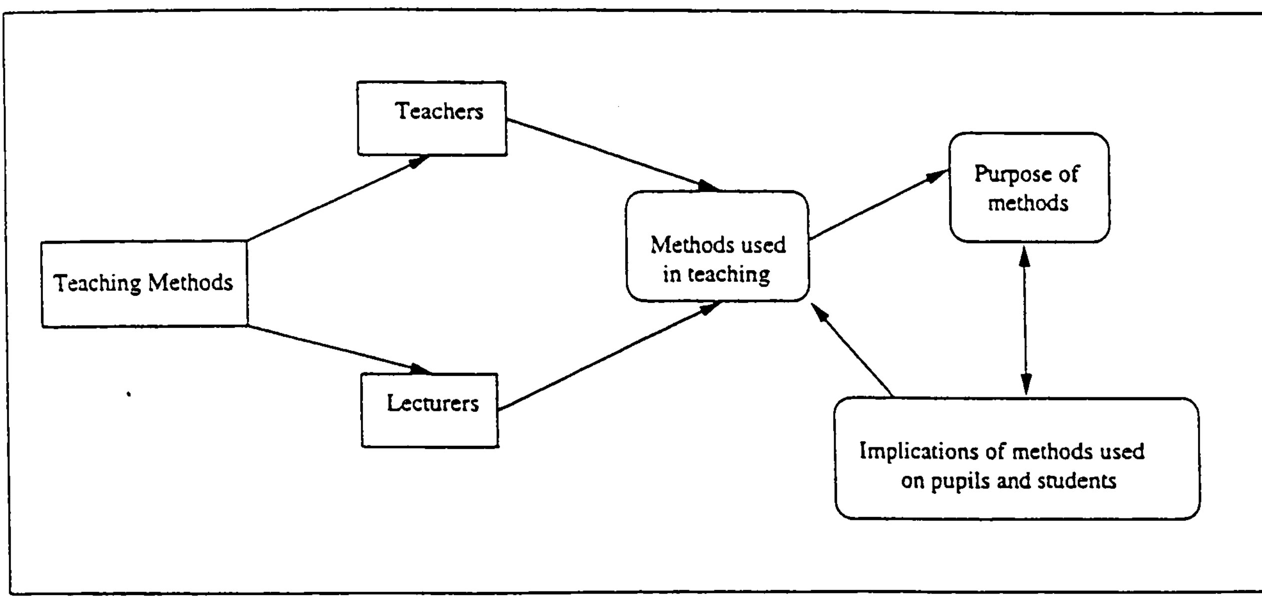
Figure 8.2(a): Conceptual Framework 1-Teaching and Learning Difficulties



Research Questions

1. What are the factors that the teachers and lecturers thought would contribute to pupils' and students' difficulties in understanding the concepts?
2. How factors that contribute to pupils' and students' difficulties in understanding concepts are associated with the areas of misconception.
3. What are the factors that contribute to teachers' and lecturers' difficulties in teaching the concepts?
4. Factors that contribute to teachers' and lecturers' difficulties in teaching the concepts are associated with the pupils' and students' areas of misconception.

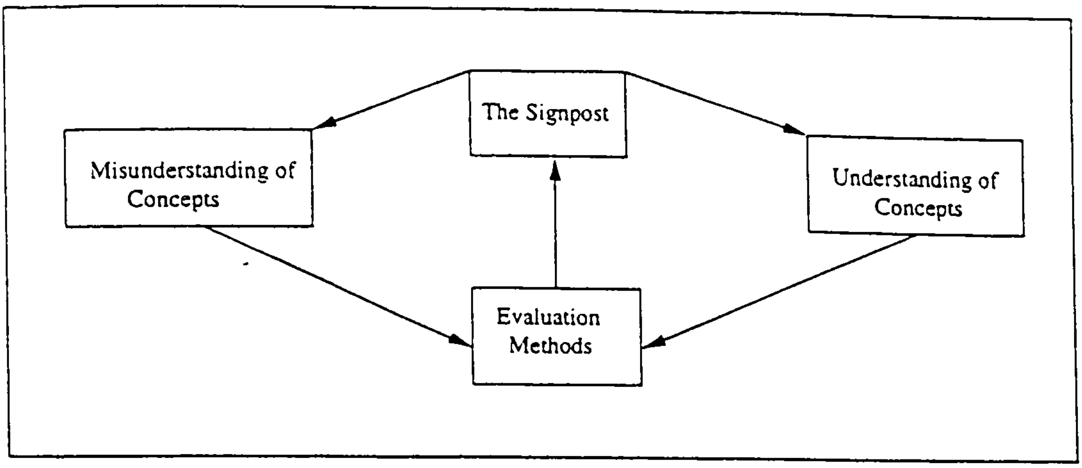
Figure 8.2(b): Conceptual Framework 2 - Teaching Methods



Research Questions

1. What are the methods that teachers and lecturers used in teaching the concepts?
2. What are the purposes of using the teaching methods?
3. What are the implications of the methods used on the pupils and students from teachers' and lecturers' perspectives?

Figure 8.2(c): Conceptual Framework 3
- Pupils' and Students' Understanding and Misunderstanding of the Concepts



Research Questions

1. What are the signposts to indicate that pupils and students have some misunderstanding of the concepts?
2. What are the signposts to indicate that pupils and students have some understanding of the concepts?
3. What are the methods used to evaluate students' misunderstanding of the concepts?
4. What are the methods used to evaluate students' understanding of concepts?

8.4 Response Themes and Codes

Based on the initial discussion of data display, this has led to the process of organising and compressing the response data in the form of matrix displays. The compressed data which come in the form of "summary phrases and quotes" were obtained as a result of data reduction. These compressed data which represent the responses from different subjects may be formed into several themes.

For example, Teacher 3 said that, *"Pupils have not been exposed to any kind of statistics syllabus before"*. Teacher 10, on the other hand said that, *"Pupils are not used to statistics, they have never done it before..."* Judging from these responses, both have a similar response theme even though the structures of the answers are different. To enable the researcher to recognise and access the themes easily, the themes had to be categorised. Accordingly, "codes" were used to represent the themes. Here, the themes can be associated with the amount of pupils' prior knowledge in statistics or any information that are closely related to it. Hence, the appropriate code for that would be "PKNOW" which stands for "Prior Knowledge". Coding in this way is another form of data reduction.

The discussion of QDA in this and in the previous sections led to the construction of Matrix Display 1 [see Appendix A8(iii)]. This is the first 28 x 5 matrix display to be constructed where the themes and codes are associated with the 28 subjects (i.e., 14 teachers and 14 lecturers) in a two-dimensional array. Based on this, multidimensional scaling and cluster analysis were developed to assist with further analysis.

8.5 Multidimensional Scaling (MDS) and Cluster Analysis

So far we have touched on the QDA techniques of data reduction, display and summary This is a subjective approach and the researcher has tried to make it more objective by providing rules to govern the data reduction and display. This is as far as most QDA goes. However, it is possible to suggest an improvement by linking QDA into multivariate analysis techniques such as MDS and cluster analysis. MDS is the process of representing, in a small dimensional space, the distances between members of a group of objects. It works by trying to find points representing the codes such that these distances between the points are related to the similarity between the codes.

Cluster analysis refers to the process of joining or fusing together two or more groups considered to be the most similar or closest to each other. Using two agglomerative methods of hierarchic classification, i.e., single linkage and complete linkage clusetering, these may be represented by inverted tree structures or dendrograms (Everitt, 1980; Everitt and Dunn, 1991). The majority of the MDS and cluster analysis begin with the calculation of a matrix of similarities or distances between items (or codes).

For example Matrix Display 1 [see Appendix A8(iii)] was converted into a 28 x 17 matrix where the rows are the subjects and the columns represent all the codes in the display. The (i,j)th element of this 28 x 17 matrix takes the value of '1' if interviewee i mentions code j, otherwise it takes the value of '0'. In this matrix display, there are

28 respondents and 17 codes ranging from PKNOW, ATT...., LEX, LCON2,...THEO, COMP and INT. This is illustrated as follows.

Subjects	PKNOW	ATT	ABLE	TF1	.....	LEX	LCON2	.....	THEO	and so on
1	1	0	0	0	.....	1	0	.....	1	and so on
2	1	0	0	0	.....	1	1	.....	1	and so on
.										
.					.....					
.										
28	1	0	1	0	.....	0	0	.....	1	and so on

The distances or similarities between the codes are measured using Splus (programmer's manual, 1993). There are many different types of similarity measures and many ways to measure the similarities between the codes. Two types of similarity measures were carefully considered. One is where the similarities are measured as the proportion of teachers who mentioned two codes. For example, Table 8.1 shows that a=8 subjects mentioned both Attitude and Ability together; b=2 subjects mentioned Ability but not Attitude; c=3 subjects mentioned Attitude but not Ability, and d=15 subjects mentioned neither Attitude nor Ability. In this case, the similarity coefficient is $\frac{a+d}{a+b+c+d}=\frac{23}{28}$. The consequence of this measure is that items which are mentioned by the majority of people will automatically have a big similarity and hence they will be clustered together. The same applies to items which are not mentioned by the majority. This provides us with information about the items which were frequently mentioned and were important to the subjects. The drawback is that it penalises codes or items that have a low recall frequency and hence tend not to cluster those items together even if the items were felt to be important jointly.

Table 8.1: Two-way association table for two codes (ABLE and ATT)

		Attitude (ATT) Yes	No	Row Total
Ability (ABLE)	Yes	8	2	10
	No	3	15	16
	Total	11	17	28

The other type of distance measure is by using the Jaccard's coefficient. This technique measures the proportion of teachers who mentioned one or other of the codes who mentioned the codes together. From Table 8.1, the coefficient is $\frac{a}{a+b+c}=\frac{8}{13}$. This proportion is attached to the codes which were not frequently mentioned but were mentioned together by most subjects. This technique on the other hand, reduces the possibility of penalising codes that have a low recall frequency but which were often present together in the interview transcript. As a result, Jaccard's coefficient is preferred and used.

Jaccard's coefficient was also used to measure the distances between the codes in cluster analysis. In complete linkage clustering, the distance between two clusters is the maximum dissimilarity between a point in one cluster and a point in another cluster. This was used in constructing the clusters which produces a compact clusters with more group structure. This tends to provide a more informative display as compared to the single linkage clustering, in which the distance between two clusters is the minimum dissimilarity between two points in the clusters (Everitt and Dunn, 1991). As a result, complete linkage clustering will be featured in the analysis. Details of how the distances were measured in Matrix Display 1 in Splus, thus producing the graphical displays for MDS and cluster analysis can be seen in Appendix A8(iii).

8.6 The Analysis of Conceptual Framework 1: Teaching and Learning Difficulties

8.6.1 Introduction

The analysis in this section refers to the 28 x 5 Matrix Display 1 [see Appendix A8(iii)]. The intention of using this type of display is to form an overall view of the interview data. In the display, column 2 contains the subjects' teaching experience which was expected to reveal some effects. Columns 3, 4 and 5 contain interview responses in the form of short summary phrases and quotes. These were gathered based on the three main research questions as shown in the respective columns. Relevant "codes" are attached to the themes of phrases and quotes.

The analysis involved looking for relationships between the codes in each research variable, the relationships between the codes across several research variables and the number of teachers and lecturers who mentioned similar themes or characteristics. Finally conclusions were drawn concurrently from the matrix displays, summary tables and graphical displays.

8.6.2 Learning Difficulties

Findings to research question (1) on *the factors that the teachers and lecturers thought would contribute to pupils' and students' difficulties in understanding the concepts* can be seen in the third column of Matrix Display 1. Looking down at the column, it can be seen that majority of the subjects with less than 20 years of experience in teaching statistics mentioned "Prior Knowledge" as the main factor that contributes to pupils' and students' difficulties in understanding concepts.

Table 8.2 shows that 11 teachers (out of 14) and 8 lecturers (out of 14) mentioned "Prior Knowledge" as a contibuting factor. It was noticed that regardless of their experience, the teachers somewhat gave rather similar explanations with regard to pupils' "Prior Knowledge" in statistics. For example, Teacher 2 (T2) quoted that, *"Pupils still think statistics is an extension of mathematics. Their main objective is to get a numerical answer rather than trying to interpret it. It is different from everything else they have ever met before and this is their first time learning statistics in a conceptual manner".* Teacher 7 (T7) also mentioned that *"students have very little statistical background. If statistical ideas were introduced at an earlier stage, pupils might find it easier to get into grips with the ideas."*

Table 8.2: The number of teachers and lecturers who mentioned the codes with regard to: (1) pupils and students' difficulties in understanding and, (2) teachers and lecturers'difficulties in teaching the concepts.

		Teachers (n=14)		Lecturers (n=14)	
Codes	Description	10 years or less (n=10)	More than 10 years (n=4)	10 years or less (n=3)	More than 10 years (n=11)
PKNOW	Prior Knowledge	9	2	2	6
ABLE	Ability	8	2	3	7
ATT	Attitude	3	2	0	4
TF1	Time Factor 1 (pupils/students)	6	1	2	2
CD	Concept Difficulty	0	0	0	2
TCF	Teaching Fault	0	1	1	2
LCON1	Lack of Confidence (pupils/students)	0	1	0	0
PRAC	Practicals	0	1	0	2
BKNOW	Backgd Knowledge	0	1	0	0

LEX	Lack of Experience	8	1	0	0
LCON2	Lack of Confidence (teachers/lecturers)	5	0	0	0
PRESS	Pressure	2	1	2	4
TF2	Time Factor 2 (teachers/lecturers)	1	1	2	4
RES	Resources	2	0	1	0
KNOW	Knowledgeable	0	0	0	1

Looking further, 10 teachers (out of 14) and 10 lecturers (out of 14) mentioned "Ability" as factors that contribute to pupils' and students' difficulties in understanding the concepts. It was also found that 5 teachers (out of 14) and 4 lecturers (out of 14) mentioned "Attitude" as the contributing factors. It seems that both teachers and lecturers felt that "Ability" is a strong contributing factor. During the interview, some of the subjects mentioned that some of the pupils and students do not have the ability to think through the ideas quickly. This was basically noticed by subjects who mentioned "Ability" as the factor. For example, T5 mentioned that, *"Some pupils are not able to think through the ideas quickly and some learned by rote"*. T7 also mentioned that, *"The weaker pupils have difficulties thinking through the ideas without some assistance from the teacher."*

The other contributing factor that was mentioned by 7 (out of 14) teachers is the "Time Factor 1." They felt that pupils need more time to come to terms with the statistical terms and concepts and to think about the problems related to the concepts. T6 quoted that, *"They need more time to let the ideas grow on them. Given more time they may have a greater awareness and a greater understanding"*. T14 also responded that, *"....with only 2 hours per week I was not able to do practicals and investigative approach that would help the pupils to understand the concepts better."* During the interview particularly with the teachers with less than 10 years experience, many of them strongly felt that they were not given enough time to teach the course. This can be seen in Table 8.2 that "Prior Knowledge", "Ability", "Attitude" and "Time Factor 1" were frequently mentioned by most subjects as compared to the other factors mentioned by the same or different subjects.

"Prior Knowledge" was frequently mentioned by teachers with less than 10 years experience in teaching. In contrary, "Prior Knowledge" was frequently mentioned by lecturers with more than 10 years experience. In the category of teaching difficulties, "Lack of Experience" (LEX) was mentioned frequently by teachers with less than 10 years experience in teaching while "Pressure" (PRESS) and "Time Factor 2" (TF2) were frequently mentioned by lecturers with more than 10 years experience in teaching.

8.6.3 Teaching Difficulties

Moving on to research question (3) which is referring to factors that contribute to teachers' and lecturers' difficulties in teaching the concepts. The findings can be seen in the fourth column of Matrix Display 1. Looking down at the column, it can be seen that "Lack of Experience" was mentioned by a majority of the teachers with less than 10 years experience. T3 quoted that, *"In the first year, I was not sure how to go about introducing each concept. I find it very difficult to be sure of doing the right thing."* Another teacher (T14) mentioned that *she finds it difficult to put ideas across to pupils and felt that her approach to teach certain concepts was not appropriate.*

"Pressure" seems to be one of the main contributing factors for the lecturers. Table 8.2 shows that 6 (out of 14) lecturers indicated that this factor has made it difficult for them to teach the concepts effectively. For example, L9 and CL14 *felt the pressure in completing the syllabus within the specified time and that the lectures had to be rushed towards the end of the course.* Other lecturers (L11 and L12) felt *the pressure of having to show the relevance of statistics in other subject areas especially when they are dealing with more difficult concepts.* Six lecturers (out of 14) also mentioned "Time Factor 2" as the factor that relates to the difficulties in teaching.

8.6.4 Areas of Misconception

Three different codes were assigned to represent the response themes in column 5 of Matrix Display 1. They are "Theory" (THEO), "Computation" (COMP) and "Interpretation" (INT). One way of looking at the responses is by displaying the data in a slightly different format. This is accomplished by constructing Table 8.3.

Table 8.3: Content analytic summary of teachers' and lecturers' views on pupils and students areas of misconception

Items of Misconceptions	Teachers' Views of Pupils' Misconceptions	Number of teachers mentioning items (n=14)	Lecturers' Views of Students' Misconceptions	Number of lecturers mentioning items (n=14)
THEO (Theory)	1. Fail to formulate the null and alternative hypothesis.	5	1. How to measure variability.	2
	2. Unable to differentiate between a one-tailed and two-tailed test and when to use it.	4	2. Applying rules of probability to problem situation-no idea about conditional prob.	4
	3. Meaning of significance level.	3	3. Cannot distinguish between different types of probability distributions.	2
	4. Meaning of critical region.	2	4. Regression and Correlation-knowing which is which.	1
	5. Notion of the degrees of freedom.	1	5. Sample mean as an estimate of population mean. Cannot relate sample to population.	2
			6. The idea of interval estimate, what it measures.	2
			7. Null and Alternative Hypothesis- cannot distinguish between one-tailed and two-tailed test.	3
			8. Fail to understand hypothesis rules-test statistics, pvalue, decision rule, significance level, critical region, Type1 & 2 errors	3
COMP (Computation)	1. Arrived at a numerical answer too quickly. Keen to end the calculation-fail to understand the situation.	4	1. Find counting difficult, numerically weak.	3
	2. Difficulties dealing with mathematical notations and algebra.	3	2. Poor in understanding formulas, proofs and translating it into appropriate arithmetical operations. Do things mechanically without the need to understand.	6
			3. Often expecting exact answers.	1
			4. Resistance to the idea of using computers to do simulations.	1
INT (Interpretation)	1. Poor at interpreting results and graphs. Not very good at explaining why .... explanation tend to be very vague.	4	1. Cannot interpret into some practical context.	2
			2. Cannot relate conclusions to results.	1
	2. No interest in interpreting results, don't even know what to do with the results.	2	3. Cannot distinguish variations between group of samples or populations.	1
			4. Unable to relate concepts from one sample to another.	1

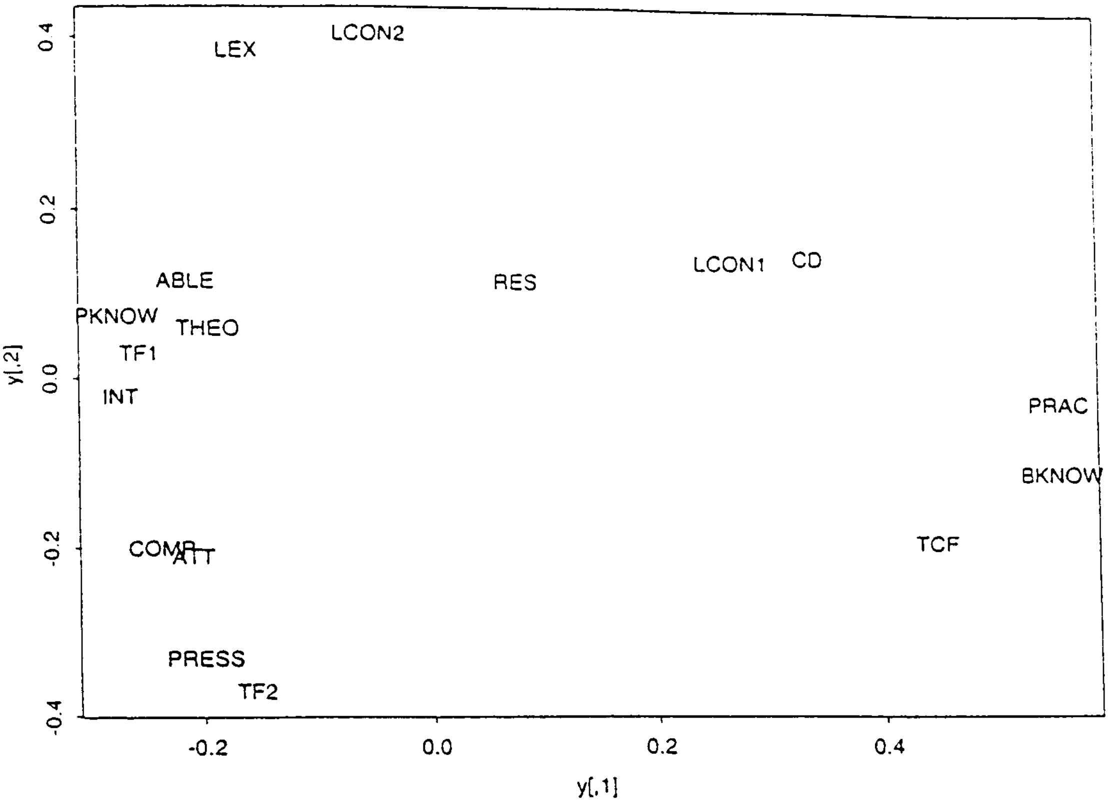
In Table 8.3, teachers' and lecturers' views of pupils/students misconception are displayed side by side for each of the codes. For the first code, i.e. "Theory", five teachers found that the pupils failed to formulate the null and alternative hypothesis and four mentioned that their pupils cannot differentiate between a one-tailed and a two-tailed test. On the other hand, the lecturers have a much wider perspective on students' areas of misconception. Four lecturers mentioned that their students were not able to apply the rules of probability to problem questions and that the conditional probability concept is greatly affected by the misconception. Three of the lecturers found that their students cannot distinguish between a one-tailed and a two-tailed test. They also failed to understand most of the procedures which are associated with the tests.

For the second code, i.e., "Computation", four teachers mentioned that their pupils were often keen to end the calculation and that they arrived at a numerical answer too quickly. They discovered that most of the pupils failed to read and fully understand the question and choose to ignore the objective of the question. Six lecturers mentioned that their students were poor at understanding formulas, and mathematical proofs and that they do things mechanically without needing to understand. Difficulties dealing with mathematical notations and algorithm were mentioned by three teachers and three lecturers. For the third code, that is, "Interpretation", four teachers and two lecturers stated that their pupils were poor at interpreting results and graphs. They felt that the pupils were not very good at explaining why and their explanation tends to be very vague.

The analysis also involved some graphical representations of the codes. This is accomplished by using the MDS technique as shown in Section 8.5. To investigate whether there is any relationship between pupils' and students' difficulties in understanding and the areas of misconception, Figure 8.3 is constructed.

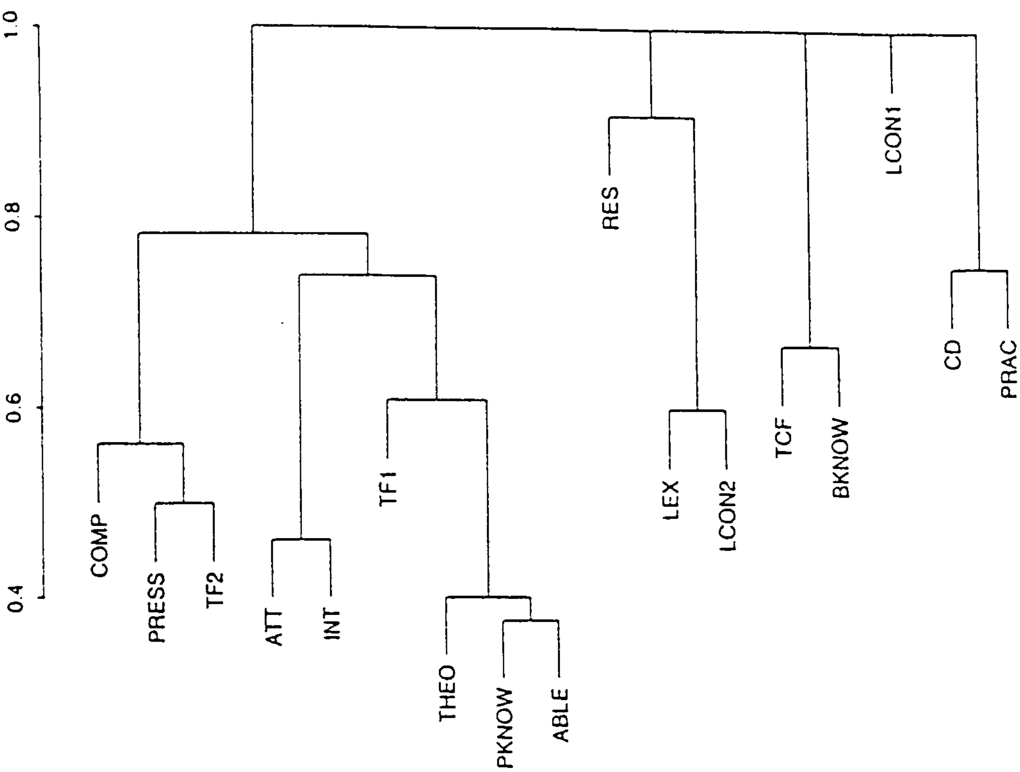
From the subjects' perspective, pupils and students who failed to understand the "Theoretical" (THEO) aspect of the subject and have difficulties in "Interpreting" (INT) the results was related to pupils and students who do not have any "Prior Knowledge" (PKNOW) in statistics, who do not acquire the "Ability" (ABLE) to think through quickly and who required more time (TF1) to understand the concepts. This is indicated by the clustering of these five codes (ABLE, PKNOW, THEO, TF1, INT) in Figure 8.3. Also, for those pupils and students who often rushed into the calculation and tended to do as quickly as they could (COMP) without needing to understand the concepts can be associated with their "Attitude" (ATT) towards the subject. This can be seen where COMP and ATT are clustered together.

Figure 8.3: 2-dimensional solution from MDS showing the relations of all response codes in Matrix Display 1



In investigating research question (4), the respective 2-dimensional solution in Figure 8.3 suggests no evidence that factors which contribute to teachers' and lecturers' difficulties in teaching the concept were associated with the pupils' and students' areas of misconception. This can be seen in Figure 8.3 where codes such as LEX (Lack of Experience), LCON2 (Teacher's Lack of Confidence), PRESS (Pressure) and TF2 (Time Factor for subjects) are quite far apart from THEO (Theory), COMP (Computation) and INT (Interpretation) codes. Further evidence of the relationship between the codes can be seen in the complete linkage dendrogram in Figure 8.4. Several clusters of codes were formed. These codes (ABLE, PKNOW, THEO), (ATT, INT) and (PRESS, TF2) which are separated by dissimilarities below 0.5 tend to form clusters which are more similar than dissimilar. Likewise, the codes which are above 0.5 tend to form pairs at relatively high distances and hence are more dissimilar. There is a resemblance in both diagrams in terms of the clustering of the codes but the MDS tends to show a more informative display as the codes which are mentioned together by the subjects could easily be identified.

Figure 8.4: A complete linkage dendrogram showing the clustering of all codes in Matrix Display 1



8.7 Analysis of Conceptual Framework 2: Teaching Methods; their Purpose and Implications

8.7.1 Teaching Methods and their Purpose

In this section, summary tables were used and constructed based on the response findings in Matrix Display 2 [see Appendix A8(iii)] which displays relevant responses from 21 subjects only. The responses from the other 7 subjects could not be elicited because of their vague responses and do not conform to the questions asked. As a result, they were excluded from the analysis. The reason for using the summary tables is because it tends to provide a more meaningful information which otherwise could not be seen clearly from Matrix Display 2. With this, more comparisons can be made between the teachers and lecturers and conclusions could be drawn easily. Table 8.4 represents the summary of the teaching methods used as well as the purpose of using the methods. Seven categories of teaching methods were formed. They are Worked Examples (WORKEX), Problem Solving through small group discussion (PSOLVE), Investigative Approach (INV), Lecturing or Teaching in a traditional way (LECT), Printed Notes and Workbooks (NOTES), Practicals and Projects (PRACS), Simulations using computers and graphic calculators (SIMS) and Others.

Table 8.4: Content analytic summary of the teaching methods and purpose of using the methods

Teaching Methods	Teachers	Lecturers
1. Worked Examples **(WORKEX)**	T2: To let pupils do questions using techniques taught by teachers. T5: To illustrate concepts in worked examples. T6: To illustrate concepts.	L8: To illustrate concepts and computations. L11: To provide theories of application for statistics in their own disciplines. L12: To explain concept and formula, what it means relative to concepts, how to work out the formula. L4: To translate a proper well-defined statistical questions for students to be able to solve.
2. Problem Solving through small group discussion **(PSOLVE)**	T4: Practice- get more experience of handling situation together. T5: To encourage them to talk about the actual assignments and investigations. T6: To encourage discussion-promote interactions between teachers and pupils. T13: To provide understanding of the back-ground.	L8: Discuss questions which they have done in advance. L11: To make them work on the problems and collaborate with one another.
3. Investigative Approach **(INV)**	T4: To relate concepts to what they are doing, say through practical situation. T6: To encourage and promote discussions- get them to think about the problem.	CL6: Students gather information and present in the form of analysis and report.
4. Lecturing/Teaching the traditional way **(LECT)**	T6: Deliver lecture and go through notes. T10: To introduce and emphasise concepts.	CL6: To present ideas and concepts. L11: To deliver materials. CL14: To introduce theory and concepts. L1: To deliver lecture materials and concepts. L7: To deliver lecture materials, talk at appropriate levels depending on who the audience is.
5. Printed Notes and Workbooks **(NOTES)**	T6: Go through notes where pupil will put in formulas, concepts and complete the examples. T10: To provide detailed information of topics.	L8: To provide more information about the topic. L9: Students filling in gaps in workbook such as formulas, examples, proofs.
6. Practicals **(PRACS)**	T13: To let them see the importance of handling real data.	L10: To impart skills as well. L9: To reinforce some of the lecture materials using Minitab. To get them to do things in the lab that might help them understand what was done in the lecture. L11: To work on Minitab- exposed them to the use of the statistical package.
7. Simulations - Computers, Graphic Calculators **(SIMS)**	T13: To build up the theory behind it, to get them to see the ideas clearly from the very basic.	L10: Looking at different variations when different samples were simulated leading to the discussion of variability, s.e., and interval estimate.

8. Others - Computer Assisted Learning (CAL) - Questions/Answers (Q/A) - Cooperative Learning - Humour, Anecdotes **(OTHERS)**	N/A	L1: To relate to practical situations. L8: Use CAL for revision purposes L11: Anecdotes - to lighten the atmosphere, to raise general awareness about the misuse of statistics. CL6: Q/A-To obtain a more precise feedback.

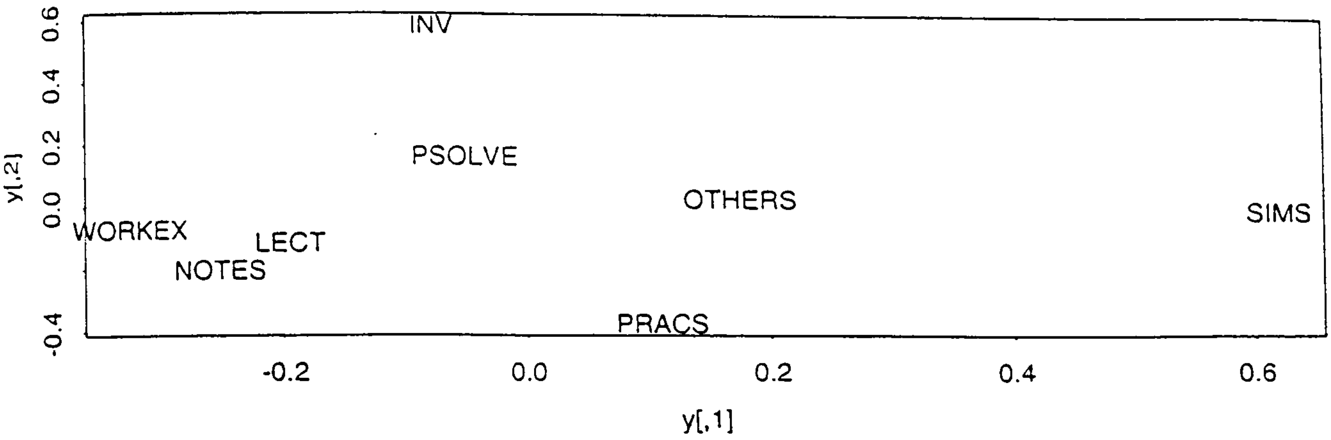
N/A: Information Not Available

All the teaching methods were used in some ways or another by the subjects. Nine subjects either used worked examples or problem solving through small group discussion or both in their teaching. The problem solving through small group discussion was used often because they wanted to encourage more discussion and interaction between the pupils/students and themselves. The other reason is because of the size of the class they were getting (ranging between 1 to 15 students per class) which enable such methods to be used.

More lecturers than the teachers mentioned the traditional lecture during the interview. Part of the reason is probably due to teachers' interpretation of traditonal teaching. Many considered it to be the normal practice at school. Hence, very few mentioned it. Only those who specifically talked about 'the chalk and talk' method were noted. Also, very few mentioned giving printed notes and workbooks to their pupils. This is because those who used the traditional teaching have already considered giving notes to their students (mainly by writing on the board during the teaching). One noticeable feature is that a majority of the teachers hardly used practicals and simulation methods or any method which involved the use of computers. This issue was raised by some of the teachers who mentioned that their schools were lacking in computer facilities and resources and thus refrained them from using the method. On the other hand, a wide variety of teaching methods were used by the lecturers in their teaching.

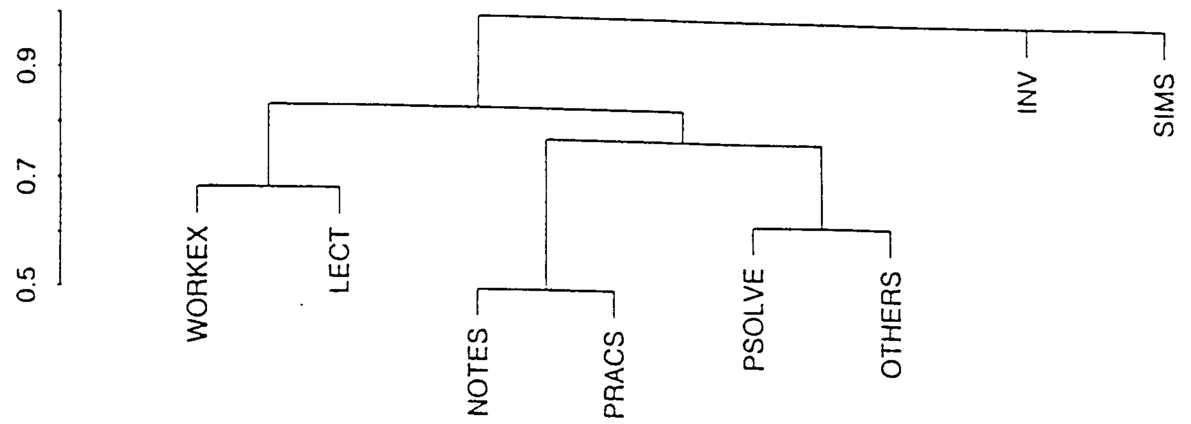
To investigate the relationships between the teaching methods used, Figure 8.5 is constructed. The diagram shows that among those subjects who mentioned one or other of these methods——Worked Examples (WORKEX), Notes (NOTES) and Lecture (LECT), these were mentioned together by the subjects. This is indicated by the items being relatively closed to each other. This implies that WORKEX, NOTES and LECT were the three methods which many subjects tend to use together in their teaching. Investigative Approach (INV) and Simulations (SIMS) methods do not appear to be mentioned together by most subjects since they are further apart from each other.

Figure 8.5: 2-dimensional solution from MDS showing the relations between the response codes on teaching methods



Further graphical display of dendrogram in Figure 8.6 shows a rather similar pattern with a formation of several clusters consisting of codes closely linked to each other. These codes (NOTES and PRACS) which forms a pair at a distance of 0.5 tends to be more similar than the codes in the other clusters [(WORKEX and LECT) and (PSOLVE and OTHERS)] which form pairs at distances of up to 0.6 and 0.7, respectively. There is a resemblance in both graphical displays as the same codes were clustered together on the left side of the displays. However, MDS display is much preferred since the distances between the codes clearly shows the extent to which the codes are mentioned together by the subjects.

Figure 8.6: A complete linkage dendrogram showing the clustering of codes on the methods used in teaching statistics



8.7.2 The Implications of the Methods Used

Table 8.5 shows the implications of the teaching methods used on the pupils and students [see also Appendix A8(iii)]. The aims are to investigate: (1) the teaching methods which were used often by the subjects, (2) the extent to which the teaching methods were used and, (3) in which circumstances the teaching methods were used. Several implication items are associated with each of the teaching methods used. For example, T2 *used worked examples in his teaching. Besides using it to illustrate and stress the concepts, he also used it to find out which particular aspect of the worked examples students found difficult to understand.* From T2's perspective, *the implication is that with his assistance pupils began to understand better (Provide Assistance). He also felt that the pupils have the opportunity to discuss with him about their difficulties in understanding some of the concepts (Promote Discussion and Interaction).* The other subject (T6) mentioned that *by using worked examples, she felt that the students eventually grasp the idea but it takes quite a while for them to be able to do so (Promote Understanding).*

On the other hand, *pupils like the examples and doing the exercise, however, they tend to rushed into the calculation without needing to understand the concept behind it.* L4 mentioned *that a lot of the students find it difficult to understand the questions and in disentangling the question story, even with the help of the worked examples.*

In the case of using the problem solving method through small group discussion, T4 felt that *the pupils have a much better working knowledge of the topic where they could get their ideas together by working on it for several hours (Promote Understanding).* In contrast, L1 felt that *the involvement of tutors in small group discussion may not solve the problem. He mentioned that the discussion process is slow, that it takes a long time for the tutors to explain to students. Also, students may take a longer time to complete certain tasks.* In general, lecturers tend to provide a wider perspective on the implications of using practicals and projects, simulations and other methods as compared to the teachers who mainly emphasised using worked examples, problem solving, investigative approach and giving printed notes.

8.8 Analysis of Conceptual Framework 3: Students' Understanding and Misunderstanding of the Concepts

8.8.1 Introduction

To assist with the analysis of Conceptual Framework 3, summary tables were constructed based on Matrix Display 3 data. This matrix display provides the information on the signpost for pupils' and students' understandings and misunderstandings of the concepts and also the methods that the subjects used in evaluating pupils' and students' understanding of the concepts. Both the Matrix Display 3 and summary Tables 8.6 to 8.8 are in Appendix A8(iii).

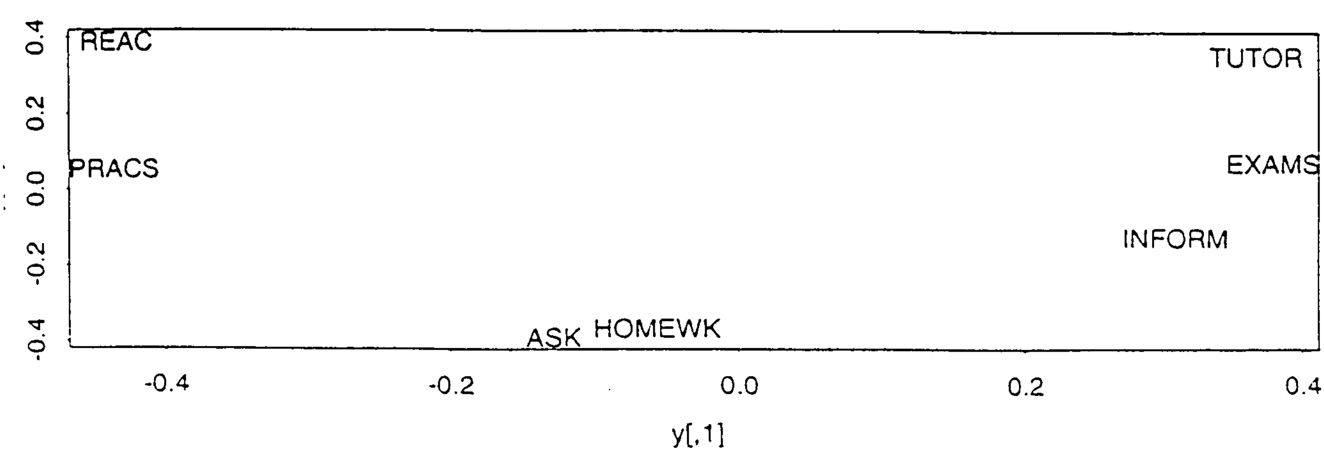
The matrix displays information on 20 subjects only. There was virtually no information from the other 8 subjects due to the way the subjects responded to questions such as "How do you evaluate your students' understanding of the concepts?" or "How do you know if your students have understood the concepts?" With these type of questions, responses tend to be subjective. One of the subjects may respond to the question such as, *"Looking at the way they present the solution to the tutorial question would give me some idea of whether they have understood the concepts or not. If they produce the correct technique and procedure, this indicates that they have understood the concept, otherwise, they have not."* In such cases, the researcher had to use her own discretion to summarise and categorise the responses accordingly. In these eight cases, categorising the responses has proved difficult due to vague responses and they were therefore excluded from the analysis.

8.8.2 Evaluation Methods

Table 8.6 display the subjects' views on the methods used to evaluate pupils' and students' understanding and misunderstanding of the concepts. The subjects' responses on the methods of evaluation were grouped together and given codes. The methods include informal evaluation such as *relying on the students to inform the teachers of their problems* (INFORM), *from the way they asked questions on certain topics* (ASK) and *by looking at their reaction such as from facial expression and body languange* (REAC). Formal evaluation includes tests and examinations (EXAMS), *homework assignments* (HOMEWK), tutorial worksheets (TUTOR) and projects, practicals and computer exercises (PRACS). Eleven (out of 20) subjects mentioned that they evaluate their students' understanding of the concepts by looking at their examination performance in which 8 (out of 12) were lecturers. The 6 (out of 20) subjects who mentioned tutorial worksheets as a method of evaluating students' understanding were also lecturers.

Further analysis of 2-dimensional solution in Figure 8.7 indicate a formation of three different clusters. The prominent linked group is found to be TUTOR, EXAMS and INFORM. This indicates that among those subjects who mentioned one or other of these codes, all three were usually mentioned together. However, the link is weak because there is very little reliable information as 7 out of 20 subjects did not have a link between 2 or more codes.

Figure 8.7: 2-dimensional solution from MDS showing the relations between the response codes on the methods used to evaluate pupils' and students' understanding/misunderstanding of concepts



8.8.3 Signposts for Pupils' and Students' Misunderstanding of the Concepts

In an interview with Teacher 2 (T2) for example, he talked about *how he thinks his pupils have very little or no understanding of what was being taught.* T2 was asked to describe further the condition which makes him realise that his pupils were not understanding the concepts. His response was *he found out from the students' homework assignment and pointed that the pupils did not use the appropriate method in testing the hypothesis.* This is in accordance with the research question——"What is the signpost for pupils' misunderstanding of the concepts?"

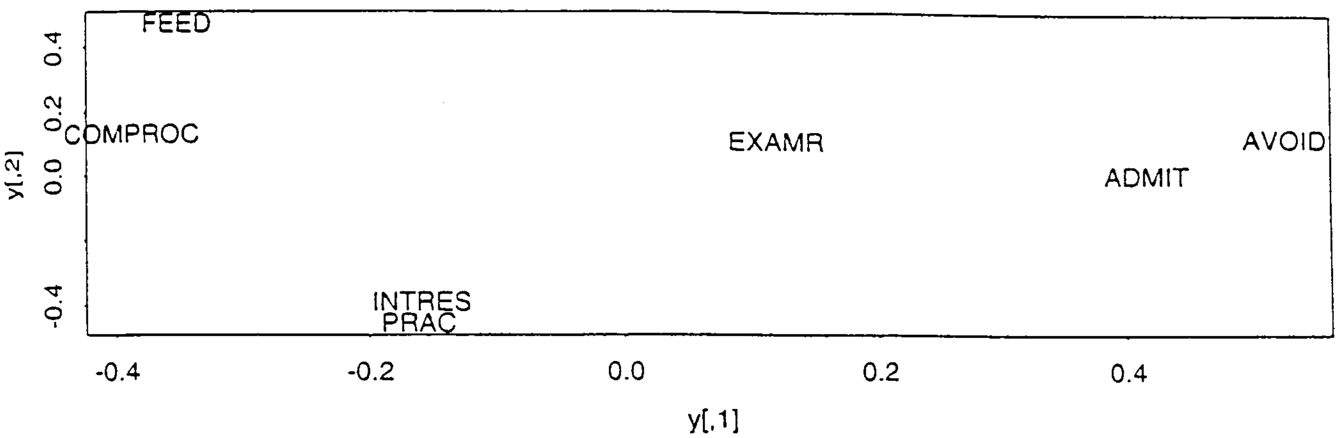
To investigate further the signposts for pupils' and students' misunderstanding of the concepts from the subjects' perspective, Table 8.7 is constructed. The findings were arranged separately for the teachers and lecturers with different codes used to represent the signpost items. These include Computational Procedures (COMPROC), Feedback on Questions Asked (FEED), Interpretation of Results (INTRES), Examination Results (EXAMR), Admitting to Problems (ADMIT), Avoiding Questions (AVOID) and Relate to Practical Situations (PRACS).

Table 8.7 shows that more teachers mentioned Interpretation of Results (INTRES), Examination Results (EXAMR) and Admitting to Problems (ADMIT) as the signposts for pupils' misunderstandings of the concepts. For example, T4 and T11 said that one of the signposts is that *pupils were unable to "interpret the results" appropriately* (INTRES). The other signpost is *when the pupils' "admitted"* (ADMIT) *to not being able to do the questions and pointed to the concept which they were having difficulties in understanding.*

On the other hand, the majority of the lecturers mentioned computational procedures (COMPROC) and feedback from the students (FEED) as the signposts for students' misunderstanding. For example, L1 realised that *the students were not understanding when they failed to use the appropriate statistical techniques in performing certain calculations.* L2 also realised that *students made very few attempts to answer questions involving meticulous calculations and proofs.* With regard to students' feedback either from their assignments (formal evaluation) or from the way they responded to the test or tutorial questions (informal evaluation), lecturers think that they mainly have a good impression of students' misunderstanding when they presented the wrong solutions to the questions.

Figure 8.8 indicates a formation of three groups of items. Two of the groups which were of interest are 'COMPROC and FEED' and 'ADMIT and AVOID.' The signpost for misunderstanding items COMPROC (Computational Procedures) and FEED (Feedback on Questions Asked) were mentioned together by the subjects who also mentioned one or other of these items (or codes). Again the links between the codes are weak as 12 (out of 20) subjects did not have a link between two or more codes.

Figure 8.8: 2-dimensional solution from MDS showing the relations between the response codes on the signposts for pupils' and students' misunderstanding of concepts



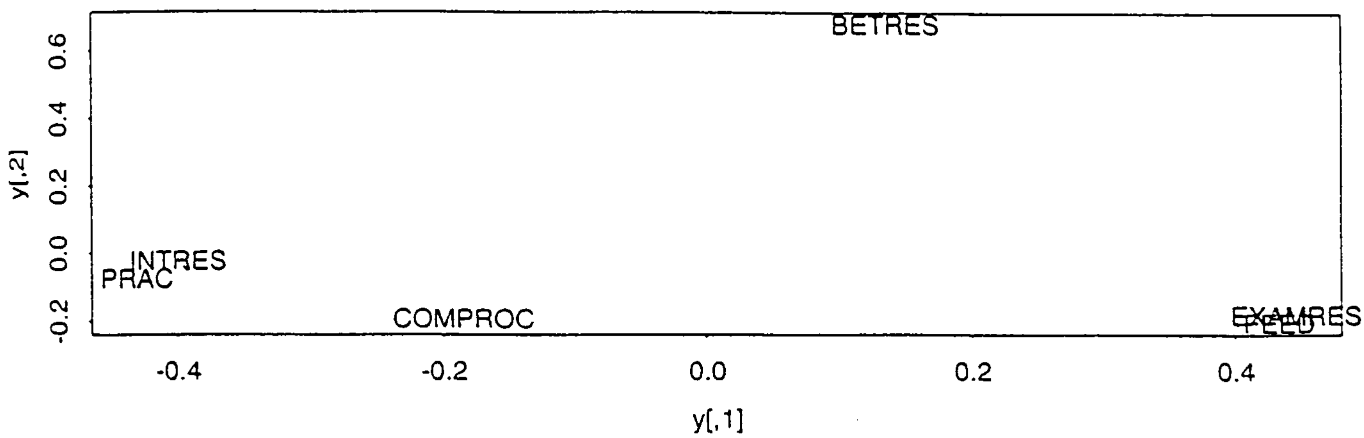
8.8.4 Signposts for Pupils' and Students' Understanding of the Concepts

From Table 8.8, it was gathered that 4 subjects felt that the students understand the concepts if they used the "appropriate statistical techniques and computations" (COMPROC) in solving the problem exercises. One of the teachers (T13) quoted that, *"Once they are able to do the work with a variety of problem exercises, you know somehow they have understood it."* Three subjects in which two of them were teachers evaluate students' understanding from the feedback they received from students written homework assignments (formal evaluation) and from the nature of students verbal answers to the questions asked (informal evaluation). For example, L8 quoted that, *"From the way they responded to the questions...if answered correctly, I can get the understanding that they have managed to grasp the idea."*

Five subjects in which four of them were teachers evaluate pupils' and students' understanding by looking at their examination performance (EXAMRES). The subjects felt that pupils and students who passed and obtained good results have achieved some understanding of the concepts. For example, T8 quoted that, *"if they pass means they do understand some of the questions. They have to understand before they can actually produced the required solutions."* Three subjects including one teacher evaluate by looking at how their pupils and students handled the practicals. They felt that pupils and students who were able to produce the appropriate results, analyse accordingly and able to make inferences and decisions were considered to have gain some understanding of the concepts.

The 2-dimensional solution in Figure 8.9 shows the linked group 'INTRES, PRAC, COMPROC' as prominent, with subjects mentioning two or more of these codes. Again, there is very little reliable information as the reliability of information is weakened by a few subjects who mentioned two or more of these codes.

Figure 8.9: 2-dimensional solution from MDS showing the relations between the response codes on the signposts for pupils' and students' understanding of concepts



Most of these findings were determined and represented by those subjects who mentioned two or more codes and have provided relevant responses to the questions. The lack of response was also attributed to the manner in which the questions were prompted and the nature of the responses produced during the interview. In the analysis of Conceptual Framework 3, the MDS graphical display was chosen over the complete linkage dendrogram as the former tends to provide a more informative display of the codes. The distances between the codes are higher when few subjects mention the codes together than when more subjects mention them together.

8.9 Highlights of the main findings

The factors that contribute to pupils' and students' difficulties in understanding the statistical concepts as reported by the subjects are mainly due to their

- lack of prior knowledge in statistics,
- lack of ability in understanding basic statistical concepts,
- attitude towards learning statistics.

The factors that contribute to difficulties in teaching as reported by the subjects are due to their

- lack of experience in teaching (particularly those with less than 10 years experience),
- lack the time to teach effectively,
- pressure to teach effectively.

The relationship between pupils' and students' difficulties in understanding and the areas of misconception indicated that pupils who have difficulties with the interpretation and the theoretical aspects of statistics were related to pupils' and students' prior knowledge in statistics and their attitude towards learning the subject. There was no evidence to suggest that pupils' and students' difficulties in understanding the concepts is related to teachers' and lecturers' difficulties in teaching. The teaching methods as reported by the subjects pointed to the use of lectures, tutorials, worked examples, notes and practicals as the main methods used in teaching statistics. These were among the methods that the subjects tended to use together in their teaching. As regards the methods used to evaluate understanding and misunderstanding of the concepts, the subjects tended to use these three methods together, i.e., tutorial exercises, examination solutions and through informal conversation to evaluate.

8.10 Conclusion and Discussion

This chapter has described the QDA, MDS and cluster analysis as techniques which can be combined in the analysis of the interview data. These methodological techniques have been used together to look for the teaching and learning factors in supporting the choice of the teaching model. The intention was also to investigate the feasibility of using these techniques for use in large scale study. This coincides with the methodological principle in Konolds' (1989) and Pfannkuch and Browns' (1996) studies where the information which they gathered from their interviews were used as basis for the design of their larger study. In the present study, the analysis has revealed several important findings but choosing appropriate methods to analyse the interview responses has proved difficult. In QDA, the difficulties were encountered during the construction of the conceptual frameworks and in building the matrix displays.

One of the advantages of QDA is that a thorough analysis of individual responses can be done. The interview responses can be transformed into a set of meaningful findings. This is assisted by the use of MDS and cluster analysis which help to further refine the analysis and produce a more informative display of the subjects' responses. Some of the analyses presented stretch the techniques discussed to the limits. Much of this is a result of two conflicting aims. QDA is going to work best with detailed information on each subject. For MDS and cluster analysis to be successful, reliable estimates of similarity are needed, which also means having a large sample of subjects. These techniques (QDA, MDS and cluster analysis) were used as complementary strands. The approach taken to link the themes/codes was also meant to illustrate the application and investigating the feasibility of using MDS within the qualitative data approach.

The possible criticisms as regards over interpretation of scaling based on too few subjects in the analysis is recognised and much could be due to random effects. Ideally, it would have been better to interview more subjects, however, such an investigation was not feasible given the constraints of time and cost. However, part of the reason for this work is to investigate if the two techniques can be successfully combined and the answer to this is yes.

In the analysis, it was found that teachers and lecturers do not differ much in their opinions on teaching and learning of statistics and methods of evaluating pupils' and students' understanding and misunderstanding of the concepts. The lecture method, notes, worked examples, small groups discussion, cooperative learning and teaching time were among the factors featured in the development of the teaching model. Regardless of the teaching roles, these factors were mentioned frequently by the subjects during the interview. It is expected that prior knowledge, ability, attitude and time factor were beyond control and these were among the factors mentioned by the subjects that makes their teaching task more difficult.

The development of the teaching model also could be supported by the main findings from the interview:

- Lecture method which includes the provision of notes and examples were mentioned by most subjects, both at school and university.
- Working together such as solving problems through small group discussion tends to be used in small classes at school and during the tutorial sessions at the university.
- Sufficient amount of time should be allocated for teaching and activities which involved students' learning. Time is one of the factors which many teachers feel has contributed to the lack of effectiveness in teaching.

The response data can be analysed using other techniques such as the interrelation of pattern codes and concept mapping of individual responses (Miles and Huberman, 1994). However, this may require the use of sophisticated Qualitative Data Analysis computer software such as NUDIST (Richards and Richards, 1994). These methods allow one to make connections between the codes (categories or information) and to develop higher-order classifications and categories of response themes.

For future work, in investigating similar aspects of teaching and learning of statistics, a questionnaire could be constructed based on the information gathered from this study. This is ideal for those teachers and lecturers who do not wish to be interviewed or want to stay anonymous in providing their responses to the above issues. In view of the economics of the situation, this approach would be feasible as such an investigation incurs less cost and time as compare to the interview technique. But the possible criticism still remains, "Can questionnaire technique ensure that adequate responses be obtained?" The answer to this lies in the extent to which both questionnaire and interview are used as complementary techniques as means of eliciting adequate information from the respondents.

Chapter 9

CONCLUSION AND DISCUSSION

9.1 Introduction

The conclusions to this thesis fall into three parts. Firstly we review statistical knowledge and the results obtained from the study of teaching statistics in Scotland. This is followed by a review of the development and testing of the teaching model. In this section we recap on the results and attempt to relate the findings to other relevant studies. The second major theme of the study, namely the comparison of a number of methodologies for educational studies, is reviewed in the third section. The final sections are used to discuss the major limitations of the work and the important directions for future studies.

9.2 Statistical Knowledge

In Scotland, the problems encountered by students in learning statistics may have stemmed from the way statistics was taught at the school level. There statistics was given a low priority and the actual amount of statistics taught at both the primary and secondary level was not known. The fact that there was also a lack of progression in teaching statistics at the upper secondary level (S4 and S5) may have led to statistics being resented by most students at the higher education level. Up until recently, little effort has been made towards bridging the gap in statistics teaching at the school level.

The implication is that pupils who intended to pursue their studies at the university level will have very little experience of statistical ideas such as probability and statistical inference, as these concepts were only introduced at the sixth year level (S6) in a course which many pupils will not take. Students who have not been exposed to these concepts at school level will find it very hard to grapple with the concepts in higher education and will tend to perceive statistics as being difficult to understand.

There is also a large demand for statistics teaching in the universities at the introductory level in Scotland. Hence, there is a need for a serious investigation of the teaching and learning of statistics because it is at this level that many students are exposed to learning statistical concepts for the first time. Indeed this study came at the right time when educators were being urged to invest more time researching into the teaching of statistics at the introductory level. The present study on the teaching of statistical concepts has successfully investigated several important aspects of statistics teaching at the introductory level in Scotland. More studies on the teaching of statistical concepts at the introductory level should be conducted as effective teaching of elementary statistical concepts will be needed for a long time.

There are more studies on teaching probability and probabilistic reasoning than on teaching statistical inference concepts such as sampling distributions, significance testing, confidence intervals, correlation and regression. The implication is that these statistical concepts are important in life and their practicality should be emphasised. The above discussions provide the background and incentive for later studies.

9.3 Review of the Development and Testing of the Teaching Model

The study on teaching statistics was inspired by the growing interest among the statistics educators in conducting research in educational statistics—on methods of teaching statistics, how to make the teaching of statistics better, and in assessing students' understanding of statistical concepts. One main aim was a major attempt to make the teaching of statistics better through the development of a teaching model. Recently, several other studies were also directed towards this effort of improving the teaching of introductory statistics and proposed some changes to be made to the introductory statistics course structure (Roiter and Petocz, 1996; Yilmaz, 1996). For example, Roiter and Petocz (1996) proposed a framework for the design and analysis of introductory statistics courses which takes into account issues such as student anxiety in learning statistics. One of their approaches included investigating which methods of teaching are suitable for this type of situation. This coincides with the approach taken in the present study in that among the considerations in the formation of the teaching model framework was the study of the students' learning patterns.

The basis for the development of the teaching model stems from research on methods of teaching. One model which included the expository method with some elements of cooperative work integrated into it was found suitable as the basis of the teaching model. Research into teaching and learning of statistical concepts has provided us with more insights into the problems that students have in learning basic statistical concepts rather than the problems that teachers have in teaching the concepts. There is a link between these two aspects of research even though sometimes they are not obvious and direct.

The development of the teaching model began with two classroom observational studies: (1) direct observations on styles of lecturing, (2) self-report questionnaire on students' learning patterns. Some of the findings from the students' learning patterns were incorporated into the development of the teaching model:

- provision of printed notes;
- teach topics in small units;
- need simple and understandable lectures;
- need more examples to stress concepts;
- read textbook and lecture notes to increase understanding;
- prefer to work with peers rather than asking the lecturers or tutors for help;

On the other hand, the results from direct observational study on teaching styles were not useful for the teaching model.

The process continued with an experimental study on teaching methods, described in Chapter 4. The experiment involved comparing two methods of delivering a lecture in a normal classroom setting. The design of the experiment involved no measure of prior achievement level in statistics but the evaluation of teaching was based on the assessment of students' understanding at the end of the lecture using a questionnaire. Several other experimental studies have used the level of prior achievement to assign students at random to treatment groups but in the present experiment randomisation of subjects was not possible due to timetabling constraints. On the other hand, if randomisation were to be used, it is likely that students might feel obliged to participate in the study, which could affect their performance. Also, prior knowledge of different arrangements might influence some students to skip the class. The drawback of not using a pre-test is inability to control the initial differences in the students and to measure differences in test performance with greater precision. As such there were variations in the experimental results possibly due to factors which are outwith control such as prior knowledge and ability of students. This was featured in the large classroom experiment. The analysis in this experiment indicate a small difference between the two groups based on the test scores at the end of teaching.

Next, a more sophisticated design was proposed, one that could measure the effects of the teaching methods and their interactions. With a large classroom experiment, provision of notes and worked examples was possible. Small cooperative work could not be implemented as it was difficult to exercise control over a large number of students within the course teaching framework. In the light of this and previous recommendations, a basic teaching model framework was formulated [see Chapter 5; Table 5.2].

The teaching model comprised three teaching components——Notes [i.e. no provision of supplementary notes; provision of supplementary notes], Pairing [working individually; working in pairs] and Concepts [teaching concepts separate from the worked examples; teaching concepts in the worked examples]. The process of testing the teaching model continued with further improvements made in the design of the experiments. Finally, a 2^3 randomised factorial block design with several factors confounded and partially confounded with the block factor was used to test the efficacy of the teaching components (or factors) and to investigate the effectiveness of the teaching models. Prior to implementation of the teaching model, the feasibility of the experimental design and the adequacy of the experimental materials were investigated through a pilot study. Teaching and testing materials were further refined for use in the main experiment. This time the participation of subjects was voluntary and a pre-test was used to control for the differences in the students.

In testing the effects of the teaching model components, three logistic regression models were used. The first model looked at the pre-test to investigate different levels of difficulty over the weeks. The second model looked at the post-test to investigate the effects of the teaching factors adjusted for the pre-test. The third model looked at the post-test to investigate the effects of the teaching factors adjusted for pre-test over the weeks. In the third, it was assumed that the pre-test had taken into account all the differences between the weeks. The experiments were conducted in two phases. Phase 1 involved the teaching of several statistical concepts for eight weeks and Phase 2 involved the teaching of several statistical concepts for four weeks, a replication of the first 4 weeks of Phase 1.

The Phase 1 findings indicated that there were some differences in students' prior knowledge associated with different topics. Confidence Interval concepts in Week (2) were found the easiest while Hypothesis Testing concepts in Week (4) were found the hardest by most students. The post-test score was also found to be related with the pre-test score, that a high post-test score was associated with a high pre-test score and vice versa. In Phase 1, the main finding was that there exists an interaction between notes, pairing and concepts (NPC) on the post-test adjusted for the pre-test score. The fact that there is interaction suggests that the combination of the teaching components is important in increasing students' learning when method of teaching the concepts in the worked examples is used. However, more evidence is required to substantiate the findings and the experiment was replicated.

In Phase 2, there were some differences in students' prior knowledge associated with different topics. The Confidence Interval concepts in Week (2) were found the easiest by most students. In contrast to Phase 1, the post-test score was not found to be associated with the pre-test score. There are nine cases of pretest score 0 with a mean posttest score of 4 (see Table 6.17; Chapter 6). A possible explanation is that these students have gained little benefit from the lecture but did benefit slightly more from the remedial class. The main finding was that the interaction between notes, pairing and concepts (NPC) on the post-test adjusted for the pre-test score was again important.

The analysis of the NPC interaction suggested that there were benefits in increasing students' learning when the method of teaching the concepts in worked examples was used in conjunction with the supplementary notes and students working in pairs. Other studies have shown that working in small groups such as the cooperative learning is beneficial (Keeler and Steinhorst, 1995; Borresen, 1990; Dees, 1991). Borresen's hypotheses is that voluntary association may be more conducive to cooperation than involuntary association and Keeler and Steinhorst discovered that students prefer to adjust their groupings in ways which they found natural and which worked for them. In these studies, even though the cooperative activities varies from one study to another, there was no indication that these activities were used in conjuction with other methods of teaching apart from the use of different types of teaching activity materials. In the present study, the groups who were exposed to working in pairs (P) only or in conjunction with the use of supplementary notes (NP) or with teaching concepts in worked examples (PC) did not benefit as much as those who were exposed to all three teaching combination, NPC. The implication is that the combination of all three teaching components was required in helping students to learn and that a combination of two components was less likely to produce a good teaching model. No single component was found to be influential on its own. However, the evidence was scanty and that tests involving more students were needed to prove that this combination was effective in the teaching of all concepts. Also the range of concepts, teachers and students needs to be increased to obtain more evidence on the robustness of the proposed teaching model.

To support the teaching model, it is necessary to gather views on the teaching and learning of statistics. Thus, a survey was carried out based on a pre-defined population of teachers and lecturers who were involved in the teaching of statistics at the introductory level in Scotland. The questionnaire survey provided information on teaching methods in statistics and areas which were considered as being difficult to teach and for students to understand. The main findings pointed out to statistical inference topics as being difficult to teach and understand with lectures and tutorials as the main methods of teaching. The interviews have revealed important findings from the responses gathered from 28 respondents, however limited the sample size maybe. In the factors that contribute to students' learning difficulties (based on the respondents' view), the main findings pointed to students'

- lack of prior knowledge in statistics,
- lack of ability in understanding basic statistical concepts, and
- attitude towards learning statistics.

In the factors that contribute to teachers' difficulties in teaching, the main findings pointed to teachers'

- lack of experience in teaching statistics (i.e., those in the category of less than 10 years),
- lack of time to teach the concepts effectively, and
- pressure in completing the syllabus on time.

The main findings on the teaching methods used by the subjects in teaching the statistical concepts indicated that worked examples, notes and lectures were among the three methods which the subjects tended to use together in their teaching.

9.4 Review of the Methodological Issues

As part of the development and testing of the teaching model, issues in the methodological aspects of research were raised. The difficulties and the benefits of the methodological techniques used in various teaching experiments were highlighted and the implications noted. The fact that teaching and learning styles in statistics could be investigated using some of the techniques shown is important, and it is also important to see if they could be implemented in the statistical teaching context.

Each part of the work went through a meticulous and exploratory process and there was always a limitation to what could be achieved. As with any other new development in research, it will be exposed to criticism but constructive criticism should provide encouragement and thus lead to improvement. The work in this thesis should not be construed as the sole guide to experimental research in teaching statistics but more for providing some ideas and information for statistical educationalists to continue research in this direction.

In view of the observational studies discussion, a teaching model simply cannot be developed using direct observational study nor can the techniques be used to investigate the effectiveness of a teaching model. Direct observational techniques would be suitable for investigating the events of teaching and in analysing teaching or learning behaviour. There is less scope for this technique to assist with the development and testing the efficacy of a teaching model and its components. If direct observation was to be considered, it could be used in recording the behaviour of the students during the experimental group work. The self-report questionnaire technique was pursued due to its practicality in obtaining information. It is appropriate for gathering opinions on lecture preference and teaching methods, for measuring students' learning and, in assisting with the formation of the teaching model framework. As such this technique was used throughout the development of the teaching model and is recommended for use in other related studies.

An attempt at a large classroom experiment showed that it was feasible to experiment on teaching but somewhat limited to what could be done in comparing the teaching methods. In the experimental findings, differences between students are greater than the effects of the teaching methods. The effect of this could have been reduced by a better experimental design. By allocating the students at random to the treatment and control groups based on students' prior test achievement in statistics, the differences in the groups could be controlled and the effects of the experimental teaching method could be measured against the control group, based on normal lecture alone.

Many educational studies compared the effectiveness of two or three methods of teaching. Very few studies in educational statistics have used sophisticated statistical techniques to measure the effect of teaching and most only implemented a simple experimental design. The work in Chapters 5 and 6 implied that it was possible to do an experiment based on a factorial block design in a small classroom setting where the researcher can exercise moderate control on the experiment. This was successfully demonstrated using a 2^3 randomised factorial block design with confounding. The contribution to knowledge is that the effectiveness of the teaching models can be investigated given a suitable experimental design and effective organisation of the experiments. Here, there is an advantage because the study used an experimental design which does not only take account the structure of teaching and measures the effect of the teaching components but also the level of difficulty in the concepts taught. The small classroom experiments have contributed greatly to the development and testing of the teaching model as a result of careful organisation of the study, which have revealed important findings.

The work in this thesis also involved investigating the feasibility of using the questionnaire and interview techniques. The study has also looked at the possibility of using Qualitative Data Analysis (QDA) in statistical education and its feasibility in interview analysis. The techniques of QDA was further extended by tackling the multidimensional scaling (MDS) technique and cluster analysis, through graphical representations obtained from MDS using Jaccard's coefficient and clustering of the codes. These techniques were combined and used together in a relatively novel manner. The contribution to knowledge is that this work has demonstrated that these techniques can be successfully combined and that they should be used in future related studies. The findings indicated that QDA is useful as a means of data reduction and data display for interview data. There is scope for MDS and cluster analysis to be used in an extended and larger studies involving more subjects. It is important to get as many interviewees as possible to respond appropriately to the interview questions so that more information can be revealed. The interview technique could also be used to assist with the construction of a structured questionnaire which aims at investigating similar aspects of research. These techniques were found feasible and is recommended for used in future related studies in educational statistics.

9.5 The Limitations

Even though an effort was made to cover in detail most aspects which are related to teaching statistical concepts in some circumstances, there are other important aspects of the study which cannot be researched thoroughly.

For example, the experiment could only be conducted at the University of Strathclyde and not at other universities in Scotland in view of access and resources available in conducting the study. It was the researcher's desire to conduct the experiment in several schools in Scotland but, the process would require the researcher to seek permission from the education authorities of several regional councils as well as from the head teacher, school teachers and pupils of statistics. The experiments would have had to be prepared carefully within each school setting in order to avoid major disruption in teaching. Even if some schools had agreed to the experiment, the topics chosen for the experiment would have had to coincide with the teaching of the topics at schools. This would have required extensive monitoring of statistics teaching at schools prior to the implementation of the experiments.

Some of the experimental work was held back because of the limited resources available. The experiment had to be designed to coincide with the teaching framework and the structure of the syllabus. Specifically, the timing should be appropriate in order to maximise students' participation in the study. Even if the experiment is carefully designed and controlled, there are factors outwith the researcher's control such as students' prior knowledge, ability and attitude. Also, in some studies the subjects are paid to participate in the experiment, but again this was not possible. A fairly narrow band of students, namely students who perceived themselves as needing extra help, were used for the small group experiment. These are an important group and one where much teaching effort needs to be focused. The results would be more convincing if more students were used and a wider range of abilities and needs.

9.6 Future Research in Teaching Statistical Concepts

The section discusses where the present research on teaching statistical concepts fits in the context of other work. It begins by listing several topics which Shaughnessy (1992) feels are important for educators and researchers to research over the next decade. The topics are:

- Development of Assessment Instruments.
- Teaching Experiments——What are the Effects of Instruction?
- What are Secondary Students' Conceptions and Misconceptions?
- Cross-Cultural Studies on Teaching Statistics.
- What are the Effects of Computer Software?

Development of Assessment Instruments

There should be some standard, reliable tools to assess students' conceptions of probability and statistics. There are tasks which have been generated by psychologists and mathematics educators, among whom were Fishbein (1975) and Falk (1983) which are available for researchers to modify and choose. However, the use of these tasks in different educational settings may result in different responses being obtained. Shaughnessy (1992) stressed that both paper-and-pencil instruments and structured, repeatable interview scripts are needed to investigate students' conceptions of probability and statistics across wide ranges of grade levels and in varieties of contexts. Even though the techniques may vary slightly depending on the research objectives and age groups, still there is a need to develop more consistency in the assessment instruments across age levels. As regards the structured interview scripts, this was one of the earlier assessment instruments used in the laboratory setting to investigate students' understanding of the basic concepts of normal distribution. This was conducted within a limited time frame. However, it would be feasible to replicate the process in a different educational context with further modification made on the interview scripts.

In the light of the above discussion, there is a possibility that the paper-and-pencil instrument used in assessing students' understanding of the concepts in the small classroom experiments could be modified and used in a larger experiment involving more participation from the students. In a group with a large number of students, students could be divided into smaller groups each of which is assigned certain tasks. For example, the first task might require each group to read and understand the concept of testing a population mean using one-sided and two-sided tests and the second task to solve two problems on concept using the appropriate statistical method. At the end of the completed tasks, students are asked to fill in a questionnaire which aims to assess their conceptual understanding and computational skills of the process involved in the task.

Teaching Experiments——What are the Effects of Instruction?

This is where the small class experiments as presented in this thesis fits into the framework of testing the effects of instruction. In testing the teaching model, one aspect which was emphasised during the experimental teaching is whether students' understanding and performance change under the influence of different methods of teaching. Using a suitable experimental design, the models were tested under several experimental conditions.

It is important that both researchers and teachers form an investigative partnership, in which the teacher is a co-researcher and the researcher is a co-teacher. Shaughnessy

(1992) stressed that, "clinical teaching experiments that carefully document changes in student's statistics and probability conceptions, beliefs, and attitudes over long periods of time are needed to obtain a clearer picture of the cognitive and affective development in probability and statistics." He added that, "both teaching experiments and more standard long-term interventions with a strong clinical methodological component are needed at all levels of instruction."

In the small classroom experiments, eight teaching models (or teaching combinations) were tested. As a result of the analysis, it was established that the three teaching combination of notes, pairing and concepts (NPC) is important in helping students to understand the statistical concepts. However, no study has been conducted to investigate the effectiveness of this teaching combination elsewhere such as in schools and other universities in Scotland. One way forward is to extend the research in a larger experiment. First, the objective of the study should be made clear to the Scottish Education Authorities for their perusal prior to conducting the study at schools. School teachers and pupils should be informed of the necessity for conducting the study on teaching statistical concepts. Similarly at the higher education level, major collaboration and interest from statistics educators are needed to continue the research in that direction.

What are Secondary Students' Conceptions and Misconceptions?

Shaughnessy pointed that there is very little large-scale information available about how secondary school students think about chance, random events and decisions under uncertainty. One aspect which needs to be investigated is that, "do students' conceptions change under the influence of instruction?" Even though Shaughnessy's comments on the lack of research in this particular area was made in view of his investigation in the United States, this problem also arises in the UK, particularly in Scotland.

In the investigation on the provision of statistics in secondary schools in Scotland (Chapter 1), the amount of statistics taught at S1 to S4 was very low. Hence, there has been little impetus to carry out research on the problems that students have in learning it as well as teachers in teaching it. But, there is a prospect of conducting more experimental studies and in-depth interviews to investigate students' level of conception and misconceptions. For example, the written and objective type of conceptual and computational questions used in the experiment can be refined further to investigate the issue.

Cross-Cultural Studies on Teaching Statistics

Shaughnessy also indicated that most of the psychological research on decision-making under uncertainty and students' conceptions of probability and statistics has been done in a few western countries such as the United States, the United Kingdom and Germany. Indeed, there is a potential for this type of research to develop in the far eastern countries such as in Malaysia, Singapore and Indonesia. Culturally, these countries resemble each other from the aspects of having similar ethnic groups, religions and social backgrounds.

The prospect is even greater for studies on teaching statistics to be conducted in one of the largest colleges in Malaysia where the researcher is currently employed. Generally, statistics courses have been taught to many students in this college for the past 20 years. Back then, the importance of statistics was recognised but only to a small group of mathematics and statistics educators. The pattern is slowly changing so that in the last 5 years, there has been a sudden interest in statistical educational studies. However, there is still a wide range of research areas and interests which have not been investigated in Malaysia, among which are described in this and other sections. Shaughnessy (1992) noted that it would be interesting to find out:

(1) What are the influences of culture on conceptions of probability and statistics,

(2) Are misconceptions of probability and statistics just artifacts of western culture, or do they appear across many cultures?

One of the prospects for future research is the extent to which the teaching model is applicable in the Malaysian culture. There is a strong possibility that there are fewer limitations in terms of students' participation as the teacher have more control on the students. Further research could be extended in a comparative studies between the two countries, Scotland and Malaysia as regards the employment of the "Notes, Pairing and Concept" teaching model. At that college, the number of students allocated for each of the classes ranges between 20 and 25 students. Hence, the prospect of conducting small classroom experiments over there is feasible.

What are the Effects of Computer Software?

Computers provide the opportunity to create a whole new learning environment for students. However, little research and evaluation has been conducted on the effects these packages have on student conceptions of probability and statistics (Biehler, 1993). Earlier, aspects of computer based learning (CBL), namely the STEPS and QUERCUS projects were discussed as regards the Teaching Learning Technology

Programme (TLTP) project. Although there is an increasing interest in the use of these materials which have been used extensively by the statistics educators involved in the projects, however, there are very few published reports on the effects these materials have on students' learning and conception.

In the light of the above, there is scope for developing suitable instrument to investigate students' learning after using the computer based learning materials. In addition to the computer based test modules provided in the CBL packages, a paper-and-pencil instrument (e.g. self-assessment questionnaire, problem solving, inter alia) could also be used at the end of the learning task to assess students' understanding of the concepts. An in-depth interview with the students should also be used to help the educators to design better learning materials and in assessing further students' understanding of the concepts.

9.7 Final Words

Any research on teaching and learning statistics can only be made possible with continuous support from the statistics educators and learners who believe that statistics teaching and learning will improve over time and effort. This study was no exception.

From the educator's perspective, any development in methods of teaching is always seen as an important issue because this would mean one step forward towards improving their teaching. However, different educators have different perceptions of the teaching methods. One may use the same type of teaching method differently from another person. Indeed, personality traits of human characteristics cannot be avoided when it comes to teaching. Some people were born with certain qualities of teaching, that regardless of the teaching methods they use, from the learners' perception, they are still regarded as good teachers.

The teaching methods used in the experiments were no different from the methods used by other educators. Some of them have used the method for many years and will continue to use it whether they think it is effective or not. Others may use the method for several reasons among them because they feel

- comfortable in using them,
- that it is appropriate for their students,
- that they have no choice but to use the method, given limited resources,

- that they are not as innovative as other people might think,
- they have limited time to be more innovative with their teaching,
- they are not motivated to find new methods of teaching,
- they lack teaching experience.

Even though it was not possible to conduct the study deeply, an effort to develop a teaching model has been partially successful. A considerable amount of effort has been made to try and blend the three methods—Notes, Pairing and Concepts together to make it better and more interesting. Also, the support given by the statistics educators in Scotland was overwhelming. There is a strong indication that many statistics educators wanted the teaching of statistics to improve and be made important and known in the society.

On the positive aspect, this study has looked at a broad range of educational facets which include:

- comparing and contrasting the methodological aspects of educational experiments in statistics,
- the conduct of large classroom experiments on teaching methods,
- the conduct of small scale intensive experiments on teaching methods,
- survey of students' opinions on methods of teaching and their learning preferences,
- survey and interviews to elicit teachers' and lecturers' opinions on teaching methods and teaching difficulties.

It is recognised that the findings reported in this study are not definitive results on experiments in teaching. But, the researcher has looked at methodological techniques in statistical education and the design and testing of investigational experiments in the study of teaching statistical concepts. The researcher has been very careful in implementing the experiments to avoid jeopardising the standard quality of educational testing. The most important thing is that the study has not destroyed student's education by experimenting on them. In fact they benefited from being in the study because of the extra help that they received in learning statistics.

REFERENCES

Adams, R.S. and **Biddle**, B.J. (1970). *Realities of Teaching: Explorations with Videotape.* Holt, Rinehart and Winston (publishers), New York.

Aitkin, M., **Anderson**, D., **Hinde**, J. (1981). Statistical Modelling of Data on Teaching Styles. *Journal of Royal Statistical Society(A)*, **144**(4):419-461.

Aitkin, M., **Anderson**, D., **Francis**, B. and **Hinde**, J. (1989). *Statistical Modelling in GLIM.* Oxford Statistical Science Series 4, Oxford Science Publications.

Amaria, R.P., **Biran**, L.A. and **Leith**, G.O.M.(1969). Individual vs. Cooperative Learning: Influence of Intelligence and Sex. *Educational Research*, **11**:95-103.

American Statistical Association (ASA) Committee on Training of Statisticians for Industry (1980). "Preparing Statisticians for Careers in Industry." American Statistician, **34**:65-75.

Anderson, C.W., **Loynes**, R.M. (1987). *The Teaching of Practical Statistics.* John Wiley and Sons Limited.

Armitage, P. (1971). Further Thoughts on Curricula. *Royal Statistical Society News and Notes, June 6-7.*

Atkinson, A.C. (1985). *Plots, Transformations and Regression.* Oxford Statistical Science Series 4, Oxford Science Publications.

Barr, G.V. (1980). Some Student Ideas on the Median and the Mode. *Teaching Statistics*, **2**(2): 38-41.

Beard, R.M., **Bligh**, D.A., **Harding**, A.G. (1978). R*esearch Into Teaching Methods In Higher Education* (4th ed.). Society for Research Into Higher Education Ltd.

Behr, M.J., **Lesh**, R., **Post**, T.R. and **Silver**, E.A. (1983). Rational Number Concepts. In R. Losh (ed.), *Acquisition of Mathematical Concepts and Processes*, pp.91- 126. New York: Academic Press.

Behr, A.L. (1988). Exploring the Lecture Method : An Empirical Study. *Studies in Higher Education*, **13**(2):189-199.

Berdie, D.R., **Anderson**, J.F. and **Niebuhr**, M.A. (1986). *Questionnaire: Design and Use (2nd ed.).* The Scarecrow Press, Inc. Methuen, N.J. & London.

Berlowitz, L. and **Levy**, B.I. (1956). Pride in group performance and group task motivation. *Journal of Abnormal and Social Psychology*, **53**:300-306.

Berg, B.L. (1989). *Qualitative Research Methods for the Social Sciences.* Boston: Allyn and Bacon (publisher).

Bibby, K. and **Davies**, N. (1995). STEPS for Learning Statistics. *Teaching Statistics*, **17**(3):107-110.

Biehler, R. (1993). Computers in Probability Education. In R. Kapadia and M. Borovcnik (eds.), *Chance Encounters: Probability in Education. A Review of Research and Pedagogical Perspectives.* Amsterdam: Kluwer.

Blackie and **Chambers**, (1989). *Mathematics In Action Series. Mathematics In Action Group.* Blackie and Sons Ltd., W & R. Chamber Ltd.

Bligh, D.A. (1970). A Pilot Experiment To Test the Relative Effectiveness of Three Kinds of Teaching Methods. *Research Into Librarianship*, **3**(15):88-93.

Bligh, D.A. (1972). *What's the Use of Lectures?* Harmondsworth: Penguin.

Bliss, J., **Monk**, M. and **Ogborn**, J. (1983). *Qualitative Data Analysis for Educational Research: A Guide to Uses of Systematic Networks.* London: Croom Helm (publisher).

Bloom, B.S. (1953). Thought Processes in Lecture and Discussion. *Journal of General Education*, **7**:160-169.

Bogdan, R. and **Biklen**, S.K. (1992). *Qualitative Research for Education: An Introduction to Theory and Methods (2nd ed.).* Boston: Allyn and Bacon.

Borresen, C.R. (1990). Success in Introductory Statistics with Small Groups. *College Teaching*, **38**(1):26-28.

Brechting, M.C. and **Hirsch**, C.R. (1977). The Effects of Small-Group Discovery Learning on Student Achievement and Attitudes in Calculus. *American Mathematics Association of Two-Year Colleges Journal*, **2**:77-82.

Brewer, J.K. (1994). Teaching Introductory Statistics to Graduate Students in the Behavioural and Social Sciences: Unique Problems and Instructional Methods. *Proceedings of the Fourth International Conference on Teaching Statistics, (ICOTS 4).* Sponsored by ISI and NIAC. Marrakech, Morocco.

Brown, G.A. (1979). *Learning from Lectures.* Nottingham, University of Nottingham.

Brown, G.A. (1980). Explaining: Studies from the Higher Education Complex. *Final Report to SSRC* . Nottingham, University of Nottingham.

Brown, G.A. and **Atkins**, M.J. (1988). *Effective Teaching in Higher Education.* London: Methuen.

Brown, G. and **Bakhtar**, M. (1988) Styles of Lecturing: A Study and its Implication. *Research Papers In Education,* **3**(2):131-153.

Brown, G.A. and **Amstrong**, S. (1984). On Explaining: Classroom Teaching Skills, In E.C. Wragg (ed.), *Classroom Teaching Skill.* Beckenham, Croom Helm.

Brown, G.A., **Bakhtar**, M. and **Youngman**, M.B. (1984). Towards a Typology of Lecturing Styles. *British Journal of Educational Psychology,* **54**:93-100.

Butler, J.A. (1992). Use of Teaching Methods Within the Lecture Format. Dorset House School of Occupational Therapy. Headington, Oxford, England. *Medical Teacher,* **14**(1):11-25.

Carpenter, T.P., **Corbitt**, M.K. and **Kepner**, H.S. Jr. (1981). What Are The Chances of Your Students Learning Probability? *Mathematics Teacher,* **74**:342-345.

Carpenter, T.P., **Lindquist**, M.M., **Mathews**, W. and **Silver**, E. (1983). Results of the Third NAEP Mathematics Assessment: Secondary School. *Mathematics Teacher,* **76**:652-659.

Chatfield, C. (1988). *Problem Solving - A Statistician's Guide.* Chapman and Hall Ltd., London.

Chervany, N.L., **Collier**, R.O. Jr., **Fienberg**, S.E., **Johnson**, P.E. and **Neter**, J. (1977). A Framework for the Development of Measurement Instruments for Evaluating the Introductory Statistics Course. *American Statistician,* **31**:17-23.

Clayton, D., **Farrands**, P., **Kennedy**, M., **Lynch**, T. and **Robinson**, P. (1994). Computer Based Instruction Systems to Support the Learning of Introductory Statistical Concepts: A Design and Development Case Study. *Proceedings of the Fourth International Conference on Teaching Statistics (ICOTS 4).* International Statistical Institute (ISI) and the National Institute of Statistics and Applied Economics. Marrakech, Morrocco.

Clark, G.M. and **Cooke**, D. (1983). *A Basic Course in Statistics (2nd ed.).* Arnold (publisher).

Clews, J.K. (1989). *A-Level Mathematics: Attitudes of Sixth Formers, Male and Female Differences, and a Case of an Experimental Teaching Method.* Master thesis. Aslib Index to Theses Abstracts, 40(1), 40-0322, p69.

Coakley, C.W. (1996). Suggestions for Your Nonparametric Statistics Course. In electronic *Journal of Statistics Education*, **4**(2).

Cochran, W.G. and **Cox**, G.M. (1957). *Experimental Designs (2nd ed.).* Wiley Series in Probability and Mathematical Statistics. John Wiley and Sons, Inc.

Cohen, L. and **Manion**, L. (1989). *Research Methods in Education (3rd. ed.).* Routledge. London, New York.

Collier, K.G. (1985). Teaching Methods in Higher Education: The Changing Scene, with Special Reference to Small-Group Work. *Higher Education Research and Development*, **4**(1):3-27.

Costin, F. (1972). Lecturing vs Other Methods of Teaching: A Review of Research. *British Journal of Educational Technology*, **1**(3):4 -31.

Cox, D.R. (1958). *Planning of Experiment.* Wiley, Inc.

Crawshaw, J. and **Chambers**, J. (1984). *A Concise Vourse in A-Level Statistics.* Stanley Thornes (publisher) Ltd.

Cumming, G. (1983). The Introductory Statistics Course: Mixed Student Groups Prefered to Streamed. *Teaching of Psychology*, **10**:34-37.

Davidson, N. (1971). The Small Group Discovery Method of Mathematics Instruction as Applied in Calculus. Ph.D. theses. *Dissertation Abstract International*, 31(11), 5927A. Madison: Wisconsin Research and Development Center for Cognitive Learning.

Davidson, R. and **Swift**, J., eds. (1988). *Proceedings of the Second International Conference on Teaching Statistics (ICOTS 2).* Sponsored by the International Statistical Institute and University of Victoria, Canada.

Davies, N. and **Radcliff**, P. (1994). Experience with Authoring Problem-Based Teaching and Learning Materials for Computer Based Learning. *Proceedings of the Fourth International Conference on Teaching Statistics (ICOTS 4).* International Statistical Institute (ISI) and the National Institute of Statistics and Applied Economics. Marrakech, Morrocco, pp.153-160.

Davies, N. (1996). *Statistics: Communicating, Learning and Technology.* Research Report, Department of Mathematics, Statistics and Operational Research, The Nottingham Trent University.

Dees, R.L. (1991). The Role of Cooperative Learning in Increasing Problem-Solving Ability in a College Remedial Course. *Journal for Research in Mathematics Education*, **22**(5):409-421.

Deutsch, M. (1949). An Experimental Study of the Effects of Co-operation and Competition Upon Group Process. *Human Relations*, **2**:199-231.

Deutsch, M. (1960). The Effects of Cooperation and Competition Upon Group Process. In D. Cartwright and A. Zander (eds.) *Group Dynamics: Research and Theory (2nd ed.)*. New York: Row, Peterson & Co.

Dietz, J. (1993). A Cooperative Learning Activity on Methods of Selecting a Sample. *The American Statistician*, **47**(2):104-108.

Duchastel, P.C. (1974). Computer Applications and Instructional Innovation: A Case Study in the Teaching of Statistics. *International Journal of Mathematics Education in Science and Technology*, **5**:607-616.

Eckert, S. (1994). Teaching Hypothesis Testing with Playing Cards: A Demonstration. In electronic *Journal of Statistics Education*, **2**(1).

Efron, B. and **Tibshirani**, R.J. (1993). *An Introduction to the Bootstrap.* New York: Chapman and Hall.

Entwistle, N.J. and **Entwistle**, D. (1970). The Relationships between Personality, Study Methods and Academic Performance. *The British Journal Of Educational Psychology,* 40.

Etezadi-Amoli, J. (1994). Statistics in Business Schools: A Content Analysis. *Proceedings of the Fourth International Conference on Teaching Statistics (ICOTS 4).* International Statistical Institute (ISI) and the National Institute of Statistics and Applied Economics. Marrakech, Morrocco.

Evans, J. St. B.T. (1989). *Bias in Human Reasoning, Causes and Consequences.* Essays in Cognitive Psychology, Lawrence Erlbaum Associates, Inc.

Everitt, B.S. (1980). *Cluster Analysis* (2nd ed.). Halsted, New York.

Everitt, B.S. and **Dunn**, G. (1991). *Applied Multivariate Data Analysis.* Edward Arnold, a member of the Hodder Headline Group (publisher).

Falk, R., **Falk**, R. and **Levin**, I. (1980). A Potential for Learning Probability in Young Children. *Educational Studies in Mathematics*, **11**:181-204.

Falk, R. (1983). Children's Choice Behaviour in Probabilistic Situations. In Grey, D.R. et al. (eds.), *Proceedings of the First International Conference on Teaching Statistics (ICOTS 1).* Teaching Statistics Trust, centre for Statistical Education, University of Sheffield, Volume 2, pp.714-726.

Falk, R. (1986). Misconceptions of Statistical Significance. *Journal of Structural Learning*, **9**:93-96.

Falk, R. (1988). Conditional Probabilities: Insights and Difficulties. In Davidson, R. and Swift, J. (eds.), *Proceedings of the 2nd International Conference on Teaching Statistics (ICOTS 2).* University of Victoria, Canada, pp. 292-7.

Fielding, N.G. and **Fielding**, J.L. (1986). *Linking Data: The Articulation of Qualitative and Quantitative Methods in Social Research.* Qualitative Research Methods Series, Volume 4. Sage Publications.

File, J. (1984). Student Learning Difficulties and Teaching Methods. *Studies in Higher Education*, 9:191-194.

Fishbein, E. (1975). *The Intuitive Sources of Probabilistic Thinking in Children.* Reidel, Dordrecht, tr. Sheppard, C.A.

Flanders, N.A. (1970). *Analyzing Teaching Behaviour.* Addison-Wesley.

Gage, N.L. and **Berliner**, D.C. (1979). *Educational Psychology (2nd ed.).* Rand McNally (publisher), Chicago.

Garfield, J. (1981). *An Investigation of Factors Influencing Student Attainment of Statistical Competence.* Ph.D. theses, University of Minnesota. Dissertation Abstracts International, 42, 5064A.

Garfield, J. and **Ahlgren**, A. (1988). Difficulties in Learning Basic Concepts in Probability and Statistics: Implications for Research. *Journal for Research in Mathematics Education*, 19(1):44-63.

Garfield, J. (1993). Teaching Statistics Using Small-Group Cooperative Learning. In electronic *Journal of Statistics Education*, **1**(1).

Garfield, J. (1994). Beyond Testing and Grading: Using Assessment to Improve Student Learning. In electronic *Journal of Statistics Education*, **2**(1).

Garfield, J. (1995). How Students Learn Statistics. *International Statistical Review*, **3**(1):25-34.

Gilchrist, W. (1982). The Role of Practicals in Statistics. *Teaching Statistics*, **4**(1):2-4.

Gibbs, I. and **Harland**, J. (1987). Approaches to Teaching in Colleges of Higher Education. *British Educational Research Journal,* **13**(2):159-173.

Glesne, C. and **Peshkin**, A. (1992). *Becoming a Qualitative Researcher: An Introduction.* New York: Longman.

Goodchild, S. (1988). School Pupils' Understanding of Average. *Teaching Statistics*, **10**(3):77-81.

Green, D.R. (1983a). A Survey of Probability Concepts in 3,000 pupils aged 11-16 years. In Grey et al. (eds.), *Proceedings of the First International Conference on Teaching Statistics*, pp. 766-783. Sheffield, UK: Teaching Statistics Trust.

Grey, D.R., **Holmes**, P., **Barnett**, P. and **Constable**, M. (1983). *Proceedings of the First Internatinal Conference on Teaching Statistics (ICOTS 1)*-Volume 1. Teaching Statistics Trust, Centre for Statistical Education, University of Sheffield, 9-13 August.

Harding, C.M., **Riley**, I.S. and **Bligh**, D.A. (1981). A Comparison of Two Teaching Methods in Mathematical Statistics. *Studies in Higher Education*, **6**(2).

Hart, A.E. (1984). How Should We Teach the Standard Deviation? *Teaching Statistics*, **6**(1):24-27.

Harre, R. and **Secord**, P.F. (1973). *The Explanation of Social Behaviour.* Totowa, N.J.: Littlefield, Adams (publisher).

Havelock, R.G., **Cox**, P., **Huberman**, A.M. and **Levinson**, N. (1983). School-University Collaboration Supporting School Improvement, Volume 4. Washington, D.C.: American University, Knowledge Transfer Institute.

Hawkins, A. (1989a). Survey of Statistics Teachers in the UK. Unpublished report, University of London Central Research Research Fund.

Hawkins, A. (1990). Training Teachers to Teach Statistics. *Proceedings of the International Statistical Institutes' Round-table Conference*, Budapest. International Statistical Institute, Voorburg.

Hawkins, A., **Jolliffe**, F. and **Glickman**, L. (1992). *Teaching Statistical Concepts.* Longman, London and New York.

Healy, M.J.R., (1984). Prospects for the Future - Where Has Statistics Failed? *Journal of Royal Statistical Society (A)*, **147**(2):368-374.

Heritage, J. (1988). Explanations as accounts: A Conversation Analytic Perspective. In C. Antaki (ed.) *Analyzing Everyday Explanations: A Casebook of Methods,* pp. 127-144. Newbury Park, CA: Sage.

Hoffer, A.R. (1982). Development of Teaching Materials at School Level. In Grey D.R. et al., (eds.), *Proceedings of the First International Conference on Teaching Statistics (ICOTS 1)*-Volume 1, p.263. Sheffield, UK: Teaching Statistics Trust.

Hogg, R.V. (1991). Statistical Education: Improvements Are Badly Needed. The *American Statistician*, **45**(4):342-343.

Holmes, P., **Kapadia**, R. and **Rubra**, G.N. (1981). *Schools Council Working Paper 69. Statistics in Schools 11-16: A Review.* Methuen Educational.

Holmes, P. (1994). Making your A-Level Statistics Teaching Practical. In Neville Davies (ed.). *Teaching and Using Statistics.* Papers presented at four workshops for teachers of statistics. Organised by the RSS Education Committee, Royal Statistical Society.

Hosmer, D.W. Jr. and **Lemeshow**, S. (1989). *Applied Logistic Regression.* John Wiley and Sons, Ltd.

Huberman, A.M. (1978). *Evaluation of Three Objectives of an Experimental Primary School: Summary of Outcomes.* Bern: Swiss Scientific Research Council.

Huberman, A.M. (1981a). *School-University Collaboration Supporting School Improvement: Volume 1. The Midwestern State Case.* Washington, DC: American University, Knowledge Transfer Institute.

Huff, D. (1954). *How to Lie with Statistics.* Norton (publisher), London.

Hunt, N. (1995). Teaching Statistical Concepts Using Spreadsheets. In Teaching Statistics supplement, *Proceedings of the Association of Statistics Lecturers in Universities (ASLU).*

International Statistical Institute (ISI) and **National Institute of Applied Economics-NIAC** (1994). Proceedings of the Fourth International Conference on Teaching Statistics (ICOTS 4), Marrakech, Morocco.

Institut Teknologi MARA (1992). *Teaching Methodology Workshop.* In association with ITM-Northern Consortium United Kingdom twinning programme.

Jimenez, A.V. and **Holmes**, P. (1994). *Students' Understanding of the Logic of Hypothesis Testing*. This paper forms part of Project ps90-0246 sponsored by the Directorate General for Scientific and Technical Research, Madrid, Spain.

Jick, T.D. (1979). Mixing Qualitative and Quantitative Methods: Triangulation in Action. *Administrative Science Quarterly*, **24**:602-611.

Johnson, L.W. (1981). Teaching Hypothesis Testing as a Six Step Process. *Teaching Statistics*, 3(2):47-49.

Johnson, D., **Johnson**, R., **Holubec**, E. and **Roy**, P. (1984). *Circles of Learning*. Association of Supervision and Curriculum Development.

Jolliffe, F.R. (1976). A Continuous Assessment Scheme for Statistics Courses for Social Scientists. *International Journal of Mathematics Education in Science and Technology*, **7**:97-103.

Jolliffe, F.R. (1991). Assessment of the Understanding of Statistical Concepts. In Vere-Jones David et al. (eds.). *Proceedings of the Third International Conference on Teaching Statistics (ICOTS 3)*. International Statistical Institute, Voorburg, , Volume 1, pp. 461-6.

Kalton, G. (1973). Problems and Possibilites with Teaching Introductory Statistics to Social Scientists. *International Journal of Mathematics Education in Science and Technology*, **4**:7-16.

Kapadia, R. (1985). A Brief Survey of Research on Probabilistic Notions. In Bell et al. (eds.), *Theory, Research and Practice in Mathematical Education*, pp. 261-265. Nottingham, UK: Shell Centre for Mathematical Education.

Keeler, C.M. and **Steinhorst**, R.K. (1994). Cooperative Learning in Statistics. *Teaching Statistics*, **16**(3):81-84.

Keeler, C.M. **and Steinhorst**, R.K. (1995). Using Small Groups to Promote Active Learning in the Introductory Statistics Course: A report from the field. In electronic *Journal of Statistics Education*, **3**(2).

Keeling, B. and **Lenz**, E.O. (1966). The Efficiency of and Attitudes Towards Two Methods of Teaching Elementary Statistics. *New Zealand Journal of Educational Studies*, **1**(2):89-101.

Keeves, J.P. (1988). *Educational Research, Methodology and Measurement: An International Handbook*. Macmillan.

Kempthorne, O. (1980). The Teaching of Statistics: Content vs. Form. *The American Statistician*, **34**:17-21.

Kendall, M.G. (1968). On the Future of Statistics - A Second Look. *Journal of Royal Statistical Society News and Notes*, June 6-7.

Kitchen, R.D. (1969). A Study on Teaching Methods. *Australian Journal of Higher Education*, **3**(3): 218-225.

Konold, C. (1989). Informal Conceptions of Probability. *Cognition and Instruction*, **3**(3):430-450.

Lambert, R., **Bullock**, R. and **Millham**, S. (1975). *The Chance of a Lifetime?* In Lambert et al. (eds.), *A manual to the Sociology of the School.* Weidenfeld and Nicolson, London.

LeCompte, M.D., **Preissle**, J. and **Tesch**, R. (1993). *Ethnography and Qualitative Design in Educational Research (2nd ed.).* New York: Academic Press.

Lesh, R. and **Landau**, M. (1983). *Acquisition of Mathematics Concepts and Processes.* Academic Press. New York, London.

Lewis, D.G. (1968). *Experimental Design in Education.* London University Press.

Lewis, J.L. and **Brewer**, J.K. (1992). Sample and Population Diagram. *Teaching Statistics*, **14**(1):10-11.

Linsey, C.W. Jr. (1988). *Teaching Students To Teach Themselves.* New York: Academic Press.

Lloyd, D.H. (1968). A Concept of Improvement of Learning Response in the Taught Lesson. *Visual Education,* **10**:23-25.

Lofland, J. and **Lofland**, L.H. (1984). *Analyzing Social Settings: A Guide to Qualitative Observation and Analysis (2nd ed.).* Belmont, CA: Wadsworth.

Lortie, D.C. (1975). *School Teachers: A Sociological Study.* Chicago: University of Chicago Press.

Loyer, M.W. (1987). Using Classroom Data to Illustrate Statistical Concepts. In Nouri E (ed.) *Proceedings of the 2nd Conference on Teaching Statistics.* New York College, Oneonta, pp. 43-73.

Mangles, T.H. (1984). Application of Micros and the Use of Computer Graphics in the Teaching of Statistical Principles. *The Professional Statistician*, **3**(7):24-7.

Marcoulides, G.A. (1990). Improving Learner Performance with Computer Based Programs. *Journal of Educational Computing Research*, **6**(2):147-155.

Maxwell, N.P. (1994). *A Coin-Flipping Exercise to Introduce the P-value.* In electronic Journal of Statistics in Education, **2**(1).

McCloskey, M., **Blythe**, S. and **Robertson**, C. (1995). Assessment of CAL Materials. *Proceedings of the Annual Conference of the Association of Statistics Lecturers in Universities*, Birmingham University.

McLeish, J. (1968a). Student Retention of Lecture Material : A Methodological Study. *Cambridge Institute of Education Bulletin*, **3**(3):2-11.

McLeish, J. (1976). The Lecture Method. In Gage, N.L. (Ed.) *The Psychology of Teaching Method*, 75th Yearbook of the National Society for the Study of Education, Chicago.

McPherson, A., **Raffe**, D. and **Robertson**, C. (1990). *Highers and Higher Education.* A Report to the Association of University Teachers (Scotland) on Post-Compulsory Schooling in Scotland and Entry to Higher Education. Centre for Education Sociology (CES).

Mercer, C.D. and **Mercer**, A.R. (1985). *Teaching Students With Learning Problems (2nd ed.).* Columbus, Ohio.

Miles, M.B. and **Huberman**, A.M. (1980). *The Realities of School Improvement Programs: Analysis of Qualitative Data.* Proposal to National Institute of Education (funded as NIE Grant).

Miles, M.B. and **Huberman**, A.M. (1984). *Qualitative Data Analysis: A Soucebook of New Methods.* Beverly Hill, CA: Sage.

Miles, M.B. and **Huberman**, A.M. (1994). *Qualitative Data Analysis-An Expanded Sourcebook (2nd ed.).* Sage Publications.

Minitab (1993). Minitab Reference Manual, Release 9 for Windows. Minitab, Inc.

Moser, C.A. and **Kalton**, G. (1992). *Survey Methods in Social Investigation (2nd ed.).* Dartmouth Publishing Company Ltd.

Nelder, J.A. and **Wedderburn**, R.W.M. (1972). Generalized Linear Models. *Journal of the Royal Statistical Society (A)*, **135**:370-384.

Numerical Algorithmic Group, NAG (1995). The Generalised Linear Interactive Modelling System, Release 4.0 Manual.

Perrot, E. (1982). *Effective Teaching - A Practical Guide to Improving your Teaching.* Longman Group Ltd.

Pereira-Mendoza, L. and **Mellor**, J. (1991). Students' Concepts of Bar Graphs - Some Preliminary Findings. In Vere-Jones David et al. (eds.), *Proceedings of the Third International Conference on Teaching Statistics (ICOTS 3)*. International Statistical Institute, Voorburg, Volume 1, pp. 150-157.

Pfannkuch, M. and **Brown**, C. (1996). Building On and Challenging Students' Intuitions About Probability: Can We Improve Undergraduate Learning. In electronic *Journal of Statistics Education*, **4**(1).

Pike, D.J. (1994). Statistics in Practice: Why is the Reality so Different from the Expectation? *Proceedings of the Fourth International Conference on Teaching Statistics (ICOTS 4)*. International Statistical Institute (ISI) and the National Institute of Statistics and Applied Economics. Marrakech, Morrocco.

Pingel, L.A. (1993). Variability - Does the Standard Deviation Always Measure It Adequately? *Teaching Statistics*, **15**(3):70-71.

Pollatsek, A., **Lima**, S. and **Well**, A.D. (1981). Concept or Computation - Students' Understanding of the Mean. *Educational Studies in Mathematics*, **12**:191-204.

Pollatsek, A., **Well**, A.D., **Konold**, C., **Hardiman**, P. and **Cobb**, G. (1987). Understanding Conditional Probabilities. *Organisational Behaviour and Human Decision Processes*, **40**:255-269.

Radke-Sharpe, N. (1991). Writing as a Component of Statistics Education. *The American Statistician*, **45**(4):292-293.

Ramsden, P. (1992). *Learning To Teach In Higher Education*. London, New York.

Richards, T. and **Richards**, L. (1994). Using Computers in Qualitative Analysis, In N. Denzin and Y. Lincoln (eds.), *Handbook of Qualitative Research*. Thousand Oaks, CA: Sage.

Ricketts, C. and **Berry**, J. (1994). Teaching Statistics through Resampling. *Teaching Statistics*, **16**(2):41-44.

Roethlisberger, F.J. and **Dickson**, W.J. (1939). *Management and the Worker*. A Research Program conducted by the Western Electric Company, Hawthorne Works, Chicago. Harvard University Press.

Roiter, K. and **Petocz**, P. (1996). Introductory Statistics Courses-A New way of Thinking. In electronic *Journal of Statistics Education*, **4**(2).

Rouncefield, M. (1994). Teaching Statistics through Practical Work. In Neville Davies (ed.), *Teaching and Using Statistics*. Papers presented at four Workshops for Teachers of Statistics. Organised by the RSS Education Committee, Royal Statistical Society.

Rouncefield, M. and **Holmes**, P. (1989). *Practical Statistics*. Centre for Statistical Education, Sheffield. Macmillan.

Royal Statistical Society (1986). "Report of a Joint Working Party of the Royal Statistical Society and the Institute of Statisticians on Statisticians: Supply and Demand." *Journal of Royal Statistical Society*, JRSS(A), **149**:122-145.

Royal Statistical Society (1994). "Abstracts Booklets." Royal Statistical Society International Conference, University of Newcastle Upon Tyne, Newcastle, September 13-16.

Royal Statistical Society (1995). "Abstracts Booklets." Royal Statistical Society International Conference: *Communicating Statistics,* University of Shropshire, Telford. September.

RSS News (1995). "Royal Statistical Society News." Volume 22, Number 10, pp.1-2.

RSS News (1995). "Royal Statistical Society News." Volume 23, Number 1, pp.1-4.

RSS News (1996). "Royal Statistical Society News." Volume 23, Number 8, pp.1-2.

RSS News (1996). "Royal Statistical Society News." Volume 23, Number 10, pp.1-2.

Ruberg, S.J. and **Mason**, R.L. (1988). Increasing Public Awareness of Statistics as a Science and a Profession - Starting in the High Schools. *The American Statistician,* **42**(3):167-170.

Schefler, W.C. (1989). *Statistics: Concepts and Applications*. The Benjamin/ Cummings Publishing Company, Inc.

Scheaffer, R.L., **Watkins**, A., **Gnanadesikan**, M. and **Witmer**, J. (1995). *Activity-Based Statistics*. A project funded by the National Science Foundation. Grant number USE-9150836.

Schoenfeld, A.H. (1985). *Understanding and Teaching the Nature of Mathematical Thinking*. Paper presented at the University of Chicago School Mathematics Project International Conference on Mathematics Education, Chicago.

Scottish Central Committee on the Curriculum, SCCC (1996). "Arrangements for Mathematics: Higher Still." Consultation Document, May. Higher Still Development Unit.

Scottish Examination Board (1990). "Scottish Certificate of Education Short Courses." A Report by Joint Working Group on Statistics, May.

Scottish Examination Board (1991). "Revised Arrangements in the Mathematics I, II, III, IV." Certificate of Sixth Years Studies in and after 1992.

Scottish Education Department (1991). "Curriculum and Assessment in Scotland National Guideline Mathematics 5-14", August. Published by Scottish Office Education (SOE).

Scottish Education Department (1993). "Effective Learning and Teaching in Scottish Secondary Schools: Mathematics." A Report by HM Inspectors of School. Published by Scottish Office Education (SOE).

Scottish in Statistics Schools (1994). A project funded by the Scottish Higher Education Funding Council (SHEFC) under the Continuing Vocational Education Scheme, August.

SEB (1990). "Scottish Examination Board Annual Report."

SEB (1991). "Scottish Examination Board Annual Report."

SEB (1994). "Scottish Examination Board Annual Report."

Senn, S. (1996). *Cross Over Trials in Clinical Approach.* Chichester, Wiley.

Serlin, R.C. and **Levin**, J.R. (1996). Two Alternative Development of the Hypergeometric Formula: Turning the Tables. In electronic *Journal of Statistics in Education*, **4**(2).

Sharan, S. (1980). Cooperative Learning in Small Groups: Recent Methods and Effects on Achievement, Attitudes and Ethnic Relations. *Review of Educational Research*, **50**:241-272.

Sharpe, N.R. (1991). Writing as a Component of Statistics Education. *The American Statistician,* **45**(4):292-293.

Shaughnessy, J.M. (1981). Misconceptions of Probability: From Systematic Errors to Systematic Experiments and Decisions. In A.P. Shulte (ed.) *Teaching Statistics and Probability.* Yearbook of the National Council of Teachers of Mathematics, pp. 90-99. Reston, VA: NCTM.

Shaughnessy, M. (1992). Research in Probability and Statistics: Reflections and Directions. In Grouws Doug (ed.) *Handbook on Research in Mathematics Education.* Macmillan.

Solomon, D., **Rosenberg**, L. and **Bezdek**, W.E. (1964). Teacher Behaviour and Student Learning. *Journal of Educational Psychology,* **55**(1):23-30.

Soper, J.B. (1983). Helping Students to Select the Appropriate Formula. *Teaching Statistics*, **5**(1):20-24.

Splus (1993). Splus Programmer's Manual. Math Soft, Inc.

Steinhorst, R.K. and **Keeler**, C.M. (1995). Developing Material for Introductory Statistics Course from a Conceptual, Active Learning Viewpoint. In electronic *Journal of Statistics Education*, **3**(3).

Stephenson, W.R. (1990). A Study on Student Reaction to the Use of Minitab in an Introductory Statistics Course. *The American Statistician*, **44**(3):231-235.

The MEW Group (1993). *Exploring Statistics.* The Mathematics in the Everyday World Group.

TLTP Catalogue (1995). " Teaching and Learning Technology Programme (Phase 1, Spring). " Innovations in Teaching and Learning Through the Power of Technology. Jointly funded by the four higher education funding bodies: HEFCE, SHEFC, HEFCW and DENI.

Tuckman, B.W. (1972). *Conducting Educational Research.* Harcourt Brace, Jovanovich, New York.

Urquhart, N.S. (1971). Nonverbal Instructional Approaches for Introductory Statistics. *American Statistician*, **25**(2):20-25.

Van Manen, M. (1977). Linking Ways of Knowing Ways of Being Practical. *Curriculum Inquiry*, **6**(3):205-228.

Vere-Jones, D., **Carlyle**, S. and **Dawkins**, B. P. (1991). *Proceedings of the Third International Conference on Teaching Statistics* (ICOTS 3). International Statistical Institute, Voorburg.

Waggoner, M. (1984). The New Technologies versus the Lecture Tradition in Higher Education: Is Change Possible? *Educational Technology*, **3**:7-12.

Wainwright, S.J. (1984). How Should We Teach the Standard Deviation? - Revisited. *Teaching Statistics*, **6**(3):86-88.

Watts, D.G. (1991). Why is Introductory Statistics Difficult to Learn? And What Can We Do to Make It Easier? *The American Statistician,* **45**(4):290-291.

Webb, N.M. (1982). Group Composition, Group Interaction, and Achievement in Cooperative Small Groups. *Journal of Educational Psychology,* **74**(4):475-484.

Weiss, N.A. and **Hassett**, M.T. (1992). *Introductory Statistics (2nd ed.).* Addison-Wesley Publishing Company.

Windeler, J. (1994). Teaching Statistics for Students of Medicine-necessity, contents and concepts. *Proceedings of the Fourth International Conference on Teaching Statistics (ICOTS 4).* International Statistical Institute (ISI) and the National Institute of Statistics and Applied Economics. Marrakech, Morrocco.

Wolcott, H.F. (1982). Differing Styles of On-Site Research, or, "If It Isn't Ethnography, What Is It?" *The Review Journal of Philosophy and Social Science*, **7**(1 & 2):154-169.

Wonnacott, T.H. and **Wonnacott**, R.J. (1972). *Introductory Statistics (2nd ed.).* John Wiley & Sons, Inc.

Woods, J.D. (1984). Lecturing: Linking Purpose and Organisation. *Improving College and University Teaching,* **31**(2):61-64.

Ward, J.N. (1956). Group-Study versus Lecture-Demonstration Method in Physical Science Instruction for General Education College Students. *Journal of Experimental Education*, **24**:197-210.

Yilmaz, M.R. (1996). The Challenge of Teaching Statistics to Non-Specialists. In electronic *Journal of Statistics Education*, **4**(1).

LIST OF APPENDICES

Appendix A3(i)	Questionnaire on Students' Learning Styles (Lecture-Assessment Questionnaire 1)
Appendix A3(ii)	Lecture Styles Forms
Appendix A3(iii)	Tutorial Session Observation Form
Appendix A4	Assessment on Students' Learning (Lecture-Assessment Questionnaire 2)
Appendix A5	Teaching Materials for Small Class Experiments (e.g., Sampling Distribution of the Means)
Appendix A7(i)	Questionnaire to Teachers of Statistics
Appendix A7(ii)	Questionnaire to Lecturers of Statistics
Appendix A8(i):	Open-ended Interview Questions
Appendix A8(ii):	A Sample of an Interview Transcript (Lecturer 7)
Appendix A8(iii):	Matrix Display 1, 2 and 3; Summary Tables 8.5-8.8; Splus commands on measure of distances between the codes and graphical displays of MDS and cluster analysis.

Appendix A3(i): Questionnaire on Students' Learning Styles

Department of Statistics and Modelling Science
53.101 Introduction to Statistics
Lecture-Assessment Questionnaire 1

Following the result of Lecture-Assessment Questionnaire 1, we now wish to investigate students' opinion on the lecture method as well as students' learning styles.

A. General

1. Course of Study (**please underline or circle**)
   Accounting Business Studies (excluding Accounting) IBML Mathematics & Statistics
   Computer Science Science Studies Others ______________
2. Gender: Male / Female

B. Students' Opinion On The Lecture Method

3. Please tick

OPINION	Always	Often	Sometimes	Seldom	Never
(a) The teacher informs students in advance the topics of their lecture.					
(b) The teacher recapitulates the previous lecture before beginning a new lecture.					
(c) The teacher stresses or summarises main points in his lecture.					
(d) The teacher outlines what he is about to teach at the beginning of a lecture.					
(e) The teacher gives students time to copy down notes.					
(f) The teacher interacts well with the students.					
(g) The teacher ask questions to clarify students' understanding of the topic.					

4. Do you like the way the lecturer delivers his material?
   (a) Yes (b) No
   Please state your reason why.
   ____
   ____

5. What do you think the lecturer should do to enable you to understand his lectures better?

OPINION	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
(a) Give more lecture notes for copying down.					
(b) Explain lecture materials clearly.					
(c) Give individual attention.					
(d) Pause for questions during his lecture.					
(e) Present content of topic in small units.					

(f) Stress or summarise main points during his lecture.					
(g) End lecture with a summary of the topic.					
(h) Make topic lecture as simple as possible.					
(i) Give detailed printed notes for each lecture.					
(j) Use more examples or case studies to stress concepts.					

C. Students' Opinion On Their Learning Styles

6. How do you understand the material covered in a lecture?

LEARNING STYLES	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
(a) By listening to the lecture only.					
(b) By listening and reading the lecture notes					
(c) By reading the textbook as well.					

(ii) Did you buy the textbook (Introductory Statistics, Weiss & Hassett)
(a) Yes (b) No
If yes, what was your opinion of it? ____________________

7. Do you find difficulties in solving the questions from the textbook?
(a) Yes Always Often Sometimes Seldom
(b) No
(c) Don't Know

8. What do you normally do when you have difficulties in solving questions from the textbook?

LEARNING STYLES	Always	Often	Sometimes	Seldom	Never
(a) Consult the tutor.					
(b) Discuss with friends.					
(c) Refer to lecture notes and textbook.					
(d) Do not do anything about it.					

9. What do you normally do whenever you have difficulties dealing with Minitab?

LEARNING STYLES	Always	Often	Sometimes	Seldom	Never
(a) Consult the tutor.					
(b) Discuss with friends.					
(c) Refer to notes and textbook.					
(d) Do not do anything about it.					

Appendix A3(ii): Lecture Styles Forms

LECTURER : ____________________
SUBJECT : ____________________
DATE : ___________
TIME : ___________

STYLE : *EXEMPLARY PRESENTER*

This type of lecturer organises lectures logically and according to a set of objectives, avoids reading the lecture notes and offering too detailed information. The lecturer uses visual aids to show processes or procedures involved; and repeats, emphasises and summarises where appropriate. They are confident, well-structured and able presenters.

DESCRIPTION	COMMENT
1. They are most likely to write down headings subheadings and brief notes rather than whole lectures.	
2. They do not have difficulty in selecting or or structuring materials for their lectures.	
3. They inform students in advance of the topics of their lectures.	
4. They recapitulate briefly the previous lecture at the beginning of a new lecture and outline the structure of the present lecture.	
5. They keep closely to the structure of their notes.	
6. They use examples and diagrams to show relationships between ideas and to show procedures or processes.	
7. They stress important points, define new terms and summarises during and at the end of their lectures.	
8. They are most likely to state the objectives of the lecture and what they expect the student to achieve at the end of the lecture.	

LECTURER : ____________________
SUBJECT : ____________________
DATE : ___________
TIME : ___________

STYLE : *VISUAL PRESENTER*

This group of lecturers structure their lectures well, but have a tendency to provide too much detail in a single lecture.

DESCRIPTION	COMMENT
1. They usually provide full notes for the students and are least likely to use only headings and brief notes.	
2. They use examples and diagrams to show statistical procedures or processes and they usually give students time to copy down complex diagrams or statistical proofs.	
3. They are confident about the material on which they are lecturing.	
4. They often define new statistical terms and use examples to illustrate concepts.	
5. They are more likely to recapitulate briefly at the beginning of each lecture.	
6. They summarise each lecture topic and link the summary to the next lecture topic.	
7. They are more likely to state the objectives of the lecture and what they expect the student to achieve at the end of the lecture.	

LECTURER : ____________________
SUBJECT : ____________________
DATE : ___________
TIME : ___________

STYLE : *AMORPHOUS PRESENTER*

These lecturers have no selected objectives and rarely structure their lectures well.

DESCRIPTION	COMMENT
1. They are least likely to state the objectives of the lecture and what they expect the students to achieve at the end of the lecture.	
2. They do recapitulate the previous lecture before beginning their lectures but often do not outline the structure of the new lecture.	
3. They do not stress or summarise main points as well as define new terms.	
4. They sometimes do not explain lecture materials clearly and rarely end lectures with a summary.	
5. They do provide lists of topics and subtitles for their students and often explain lecture materials verbally without writing.	
6. They are less likely to use examples to illustrate statistical concepts.	
7. They fail to give students time to copy down notes, diagrams and related examples.	

LECTURER : ____________________
SUBJECT : ____________________
DATE : ___________
TIME : ___________

STYLE : ***SELF-DOUBT***

This group of lecturers have difficulty in selecting and structuring material for their lectures. They do not keep close to the content and structure of their notes and at the end of most lectures they do not feel they have achieved their objectives.

DESCRIPTION	COMMENT
1. They do not outline the objectives of the lecture.	
2. They are not able to select and deliver materials within a specified time period.	
3. They tend to write down headings and brief notes rather than detailed notes.	
4. They do inform the students in advance of the topic of the lecture and are more likely to tell students on which topic they will be examined.	
5. They feel less confident about their knowledge of the topics of their lectures and are more likely to digress from the content of their notes.	
6. They tend to use diagrams to show statistical procedures.	
7. They do not recapitulate the previous lecture before beginning a new lecture, but they do outline the topic of a lecture at the beginning.	

Appendix A3(iii): Tutorial Session Observation Form

LECTURER : ____________________
SUBJECT : ____________________
DATE : ________________
TIME : ________________

INQUIRY ITEMS	GENERAL DESCRIPTION
1. HOW IS THE CLASS CONDUCTED? - In general.	
2. ARE THERE ANY SMALL GROUP TASKS? - Describe the task.	
3. IS THERE ANY SMALL GROUP DISCUSSION? - Describe the situation.	
4. ARE THERE ANY QUESTION AND ANSWER SESSION? - Describe the response from the students. (Active, Neutral, Passive)	
5. ARE THERE ANY BRAINSTORMING SESSIONS? - If so, describe the situation.	
6. ARE THERE ANY BUZZ GROUPS? - If so, describe the situation.	
7. ARE THERE ANY INTERACTION BETWEEN THE LECTURER AND STUDENTS? - Describe the situation.	
8. DESCRIBE THE TUTORIAL MATERIALS GIVEN TO THE STUDENTS. - Are they to be handed in? - If so, when?	
9. STUDENTS REACTION IN THE TUTORIAL SESSION? - Describe their attitude (alert, bored, interesting, enthusiastic) - Do students ask a lot of questions regarding the tutorial handouts?	
10. OTHER OBSERVABLE FEATURES.	

Appendix A4: Assessment on Students' Learning

UNIVERSITY OF STRATHCLYDE
Department of Statistics and Modelling Science
53.101 Introduction to Statistics

Lecture-Assessment Questionnaire 2

This questionnaire is designed to assess understanding of the materials in todays' lecture. We wish to investigate how different teaching styles influence students' understanding of the concepts.

A. GENERAL

1. Course of Study : ....

2. Sex : Male / Female

B. CONCEPTS

3. List any key concepts you can remember about todays' lecture?

....

....

Please try to answer the following questions. You may refer to your notes if you wish.

C. COURSE MATERIALS
Indicate with a TRUE or FALSE by circling the response.

4. The null hypothesis is H_0: m = 1000 ml and H_a: m < 1000 ml. The sample mean content, $\bar{x}$, of the 30 bottles obtained is 996.8 ml with a standard deviation of 33.3 ml. Assuming the null hypothesis, m=1000 ml, is true, the probability of getting a value for $\bar{x}$ of 996.8 ml or less is 0.298. This is known as the p-value. **T or F**

5. In a Z-test of H_0 : m = 10 against H_a : m > 10, the calculated test statistic gave a p-value of 0.086. This means that the null hypothesis would be accepted if the significance level (a) is equal to 0.10. **T or F**

6. The decision rule for a hypothesis test using P-value is as follows:
If the P-value is less than the significance level, a, reject H_0; otherwise, do not reject H_0.
T or F

The U.S. Federal Highway Administration reports that the average passenger vehicle was driven 8.9 thousand miles in 1982. A random sample of $n = 500$ passenger vehicles gave an average of $\bar{x} = 8.7$ thousand miles driven for last year, with a standard deviation of $s = 6.0$ thousand miles. Let m denote the true mean distance driven per passenger vehicle for last year. If a hypothesis test is performed such that : H_0 : m = 8.9 thousand miles

H_a : m < 8.9 thousand miles

and the test statistics is -0.745,

Which of the following statements are true?

7. Reject the null hypothesis (H_0) at the 1% level of significance, when the critical value is -2.33.

T or F

8. Accept the null hypothesis (H_0) at the 5% level of significance, when the critical value is -1.645.

T or F

9. Test statistic value does not lie in the critical region at 5%, when the critical value is -1.645.

T or F

10. We conclude that the true mean distance driven per passenger vehicle for last year is not 8.9 thousand miles. **T or F**

11. No decision can be made as p-value is not given.

T or F

12. Which of the following apply to Type I error? ***(Please circle any correct statement)***

A. The mistake of rejecting the null hypothesis when in fact, the null hypothesis is true.

B. The mistake of not rejecting the null hypothesis when in fact, the null hypothesis is false.

C. The mistake of rejecting the null hypothesis when in fact, the null hypothesis is false.

D. The mistake of not rejecting the null hypothesis when in fact, the null hypothesis is true.

D. LECTURE AND COMMENT

13. How do you rate todays' lecture? **(Please circle)**

Very Poor   Poor   Fair   Good   Very Good

14. Please write down any other comments you have to make about the class, course or lecturer.

Appendix A5: Teaching Materials for Small Class Experiments

(e.g., Sampling Distribution of the Means)

Item No	Items
(1)	Pre-test Questions
(2)	Key Concepts and Outcomes
(3)	Flow Diagram of the Concept
(4) & (5)	Brief Teaching Notes
(6) & (7)	Worked Examples
(8)	Problem Exercises
(9)	Post-test Questions

Name: ____________________ (1)

CONCEPT: Sampling Distribution of the Mean ($\overline{X}$)

1. Given a population with mean μ, standard deviation σ and a random sample size n. Then, the standard error of the distribution of the sample mean ($\overline{X}$) is

A. σ B. $\frac{\sigma}{\sqrt{n}}$ C. σ^2

D. $\frac{\sigma^2}{\sqrt{n}}$ E. $\frac{\sigma}{n}$

2. When using simple random sampling from a population with mean μ , the expected value of the sample mean is equal to

A. the sample mean, $\overline{X}$
B. the population mean, μ
C. the standard error of the sample mean, $\frac{\sigma}{\sqrt{n}}$
D. the variance of the population, σ^2
E. zero.

For Questions 3 and 4, refer to the following information.

Consider a normal population with $\mu = 50$ and $\sigma = 10$. A simple random sample of size n= 25 will be used to provide a sample mean $\overline{X}$ to estimate μ.

3. What is the mean value for the sampling distribution of $\overline{X}$?

A. 2 . B. 10 C. 50 D. 25 E. none.

4. What is the standard error of the sample mean, $\overline{X}$?

A. 10 B. 2 C. 0.4 D. 5 E. 2.5

5. An automatic machine used to fill cans of soup has the following characteristics; $\mu = 15.9$ ounces and $\sigma = 0.5$ ounces. What is the probability of finding a sample of 49 cans with a mean $\overline{X}$ greater than 16 ounces, i.e. $P(\overline{X} > 16)$?

A. 0.5012 B. 0.4321 C. 0.4192 D. 0.5621 E. 0.4115

CONCEPT: Sampling Distribution of the Mean ($\overline{X}$)

Key Concepts

- Sampling distribution of the mean
- Mean of the Sample Mean
- Standard Error of the Sample Mean

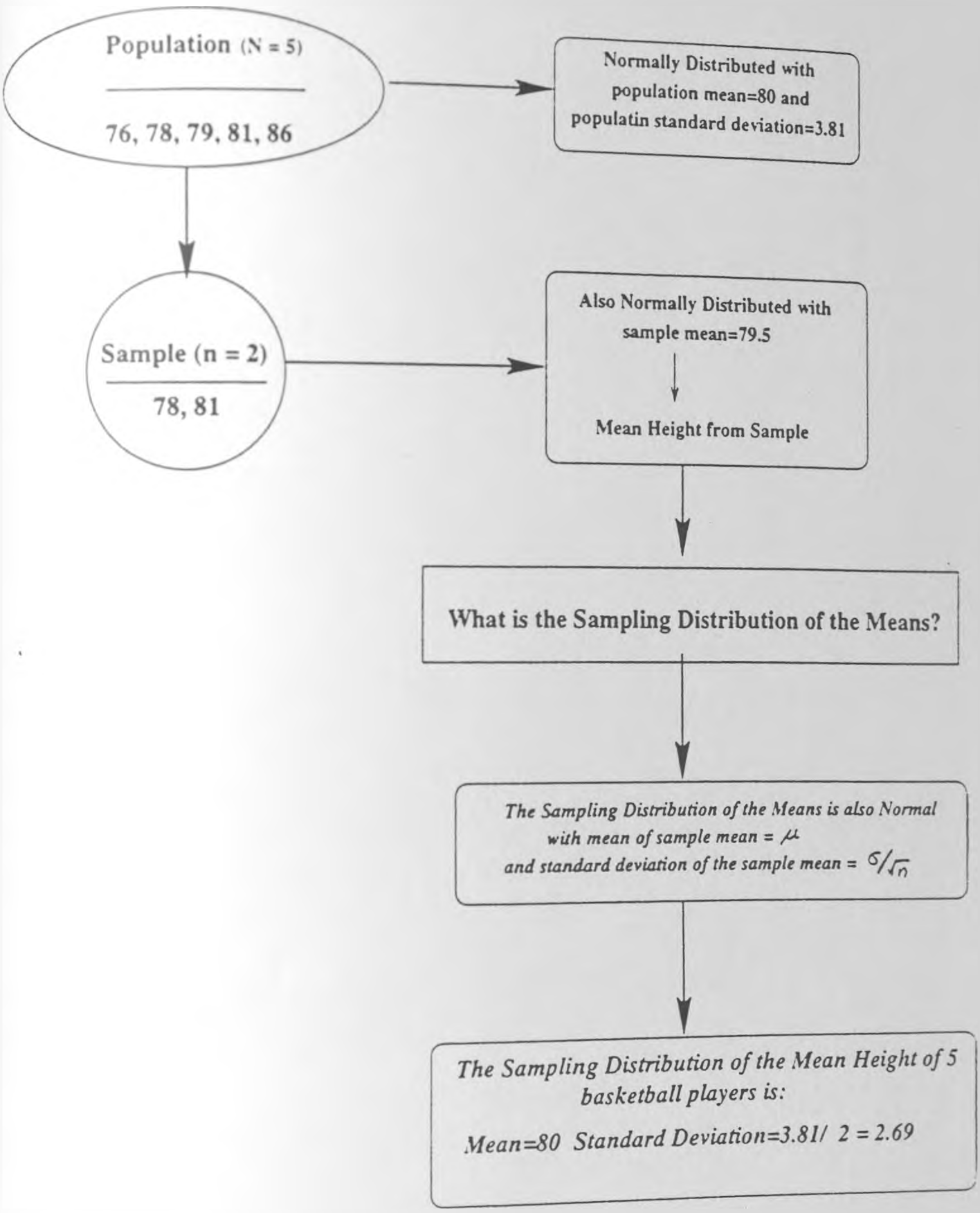
Outcomes

1. Know the distribution of the sample mean ($\overline{X}$) from a normal population.
2. Know that the mean of all possible $\overline{X}$ values is equal to the population mean μ, that is, $E(\overline{X}) = \mu$.
3. Know that the standard error of the sample mean is $\sigma_{\bar{x}} = \sigma/\sqrt{n}$

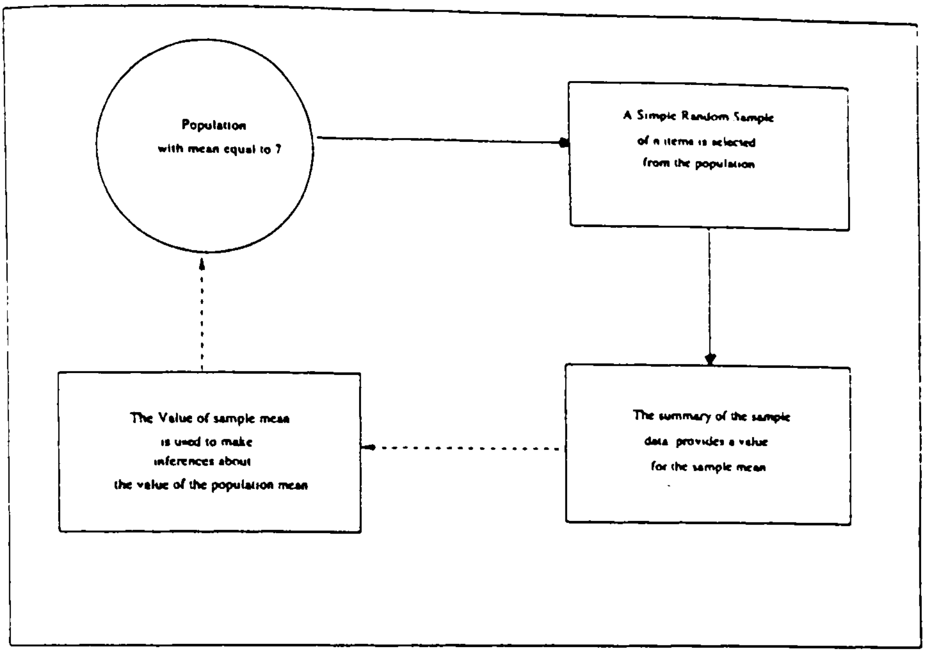
   where σ is the population standard deviation and n is the sample size.
4. Know how to calculate the probability that the sample mean lies in a given range using Normal Probability Tables, i.e. $P(a < \overline{X} < b)$.

A GENERAL PLAN TO INTRODUCE SAMPLING DISTRIBUTION OF THE MEAN

Consider a Population of Heights of 5 Basketball Players



SAMPLING DISTRIBUTION OF THE MEAN DIAGRAM



THE STATISTICAL PROCESS OF USING A SAMPLE MEAN
TO MAKE INFERENCES ABOUT A POPULATION MEAN

The sample mean ($\overline{X}$) is an unbiased estimate of the population mean (μ) -or, stated in another way, the mean for all possible $\overline{X}$ values is the same as the mean of the population from which the samples are taken,

that is, $E(\overline{X}) = \mu$

or, $\mu_{\bar{x}} = \mu$

where;

$E(\overline{X})$ denote the expected value of $\overline{X}$

or the mean of all possible $\overline{X}$ values.

KEY FACT

The sampling distribution of the mean - normal populations

Suppose a random sample of size n is to be taken from a normally distributed population with mean μ and standard deviation σ.

Then the random variable $\overline{X}$ is also normally distributed and has mean $\mu_{\bar{x}} = \mu$ and standard deviation $\sigma_{\bar{x}} = \sigma/\sqrt{n}$.

In other words, if the population being sampled is normally distributed, then probabilites for $\overline{X}$ are equal to areas under the normal curve with parameters μ and $\sigma/\sqrt{n}$.

WORKED EXAMPLES

1. Assume we have a normally distributed population with mean $\mu = 32$ and standard deviation $\sigma = 5$. Furthermore, assume that the population has 500 items and that a simple random sample of 25 items used to obtain information about this population. Let $\overline{X}$ denote the sample mean that will be used to estimate the value of the population mean.

a) What is the mean value of $\overline{X}$?

b) What is the standard deviation of $\overline{X}$?

Solution:

Given that the population is Normal with mean, $\mu = 32$

and standard deviation, $\sigma = 5$

and sample size, n = 5.

a) Refering to KEY FACT,

$$E(\overline{X}) = \mu$$

that is $E(\overline{X}) = 32$

b) Refering to KEY FACT,

$$\sigma_{\bar{x}} = \sigma / \sqrt{n}$$

that is $\sigma_{\bar{x}} = 5 / \sqrt{25} = 1$

2. The heights of a particular species of plant follow a normal distribution with mean μ = 21cm and standard deviation $\sigma = \sqrt{90}$ cm. A random sample (n) of 10 plants is taken and the mean height ($\overline{X}$) calculated. Find the probability that this sample mean ($\overline{X}$) lies between 18 cm and 27 cm, that is, $P(18 < \overline{X} < 27)$.

Solution:

Given that the heights of a particular species of a plant is normally distributed with $\mu = 21$cm and $\sigma = \sqrt{90}$ cm.

$\overline{X}$ is the sample mean height of n = 10 plants.

In order to find the $P(18 < \overline{X} < 27)$, we need to know the distribution of the sample mean.

That is,

$$\boxed{\mu_{\bar{x}} = 21 \text{ and } \sigma_{\bar{x}} = \sigma/\sqrt{n} = \sqrt{90/10}}$$

Then, we can proceed with finding the $P(18 < \overline{X} < 27)$

Standardizing, $Z = \dfrac{\overline{X} - 21}{3}$

$$= P\left(\frac{18-21}{3} < Z < \frac{27-21}{3}\right)$$

$$= P(-1 < Z < 2)$$

$$= P(Z < 2) - P(Z < -1)$$

$$= P(Z < 2) - [1 - P(Z < 1)]$$

$$= 0.9772 - [1 - 0.8413]$$

$$= \underline{0.8185}$$

Therefore, the probability that the mean height of the sample lies 18 cm and 27 cm is 0.8185.

EXERCISES

1. Weights for males between the ages of 20 and 30 have a mean $\mu = 170$ pounds with a standard deviation of $\sigma = 28$. If a simple random sample of 40 males in this age group is to be selected and the sample mean weight $\bar{x}$ computed, what are the values of mean and the standard deviation of the sample mean?

2. An automatic machine used to fill cans of soup has the following characteristics: $\mu = 15.9$ ounces, $\sigma = 0.5$ ounces.
   a. Show the sampling distribution of $\bar{x}$, where $\bar{x}$ is the sample mean for 40 cans selected randomly by a quality control inspector.
   b. What is the probability of finding a sample of 40 cans with a mean $\bar{x}$ greater than 16 ounces?

3. A library checks out an average of 320 books per day, with a standard deviation of 75 books. Consider a sample of 30 days of operation, with $\bar{x}$ being the sample mean number of books checked out per day.
   a. Show the sampling distribution of $\bar{x}$.
   b. What is the standard error?
   c. What is the probability that the sample mean for the 30 days will be between 300 and 340 books?
   d. What is the probability the sample mean will show more than 325 books checked out per day?

Name: ____________________

(9)

CONCEPT: **Sampling Distribution of the Mean ($\overline{X}$)**

1. Given a population with mean μ, standard deviation σ and a random sample size n. Then, the standard error of the distribution of the sample mean ($\overline{X}$) is

A. σ B. $\frac{\sigma}{\sqrt{n}}$ C. σ^2

D. $\frac{\sigma^2}{\sqrt{n}}$ E. $\frac{\sigma}{n}$

2. When using simple random sampling from a population with mean μ , the expected value of the sample mean is equal to

A. the sample mean, $\overline{X}$
B. the population mean, μ
C. the standard error of the sample mean, $\frac{\sigma}{\sqrt{n}}$
D. the variance of the population, σ^2
E. zero.

For Questions 3 and 4, refer to the following information.

Consider a normal population with $\mu = 70$ and $\sigma = 15$. A simple random sample of size n= 30 will be used to provide a sample mean $\overline{X}$ to estimate μ.

3. What is the mean value for the sampling distribution of $\overline{X}$?

A. 30 B. 70 C. 15 D. 5 E. none.

4. What is the standard error of the sample mean, $\overline{X}$?

A. 10 B. 4 C. 2.5 D. 2.7 E. 5

5. An automatic machine used to fill cans of soup has the following characteristics; $\mu = 15.9$ ounces and $\sigma = 0.5$ ounces. What is the probability of finding a sample of 49 cans with a mean $\overline{X}$ greater than 16 ounces, i.e. $P(\overline{X} > 16)$?

A. 0.5012 B. 0.4321 C. 0.4192 D. 0.5621 E. 0.4115

Appendix A7(i): Questionnaire to Teachers of Statistics

UNIVERSITY OF STRATHCLYDE
Department of Statistics and Modelling Science
Livingstone Tower
26 Richmond Street
Glasgow G1 1XH

From : Zamalia Mahmud
PhD Research Candidate

Supervisor: Dr. Chris Robertson

To : Teachers of Statistics

Date : 9 December, 1994

Re : **A Study on Statistical Teaching Concepts**

I am currently engaged in a study on methods of teaching statistical concepts at the introductory level. Part of the study is to look at the effectiveness of various teaching methods used in teaching CSYS Statistics in schools, the difficulties in teaching the subject, why students find certain concepts difficult and many other enquiries which led to the teaching of the subject.

The main purpose of the study is to find out that the use of various teaching methods in teaching statistics can contribute to students' understanding of certain difficult statistical concepts and that certain methods of teaching might help to diagnose the problems that students are facing in learning the subject. At a later stage, the researcher hopes to develop a statistics teaching model that could be of use in the teaching of the subject effectively. However, much investigation and co-operation are needed from teachers and students before firm conclusions can be made. The researcher feels that statistics teachers would provide a better insights into the teaching of the subject in schools and that their contribution could prove to be beneficial once the outcome of the study is exposed. In any case, please disregard questions that you feel might have offend you and that could affect your integrity as a teacher.

We intend to follow up with a personal interview to enquire more about the problems that students are facing in understanding basic statistical concepts. Hence, we would appreciate it if you could leave your name and contact number on the questionnaire and return it to us on the prepaid envelope by *15 January, 1995*. On the other hand, please let us know if you do not wish to be contacted in the future. We will assure you that all information given by those who are directly or indirectly involved in the study will be treated with utmost confidentiality. We wish to convey our thanks beforehand and would welcome any comments and suggestions regarding the study and the questionnaire.

Please answer the following questions to the best of your knowledge. All information given in the questionnaire will be treated with utmost confidentiality.

A. GENERAL

1. Name of teacher: ____________________

2. Name of School and Department: ____________________

3. Basic degree of qualification (specify area of study): ____________________
Other qualifications: ____________________
Courses and workshops attended on teaching statistics:
____
____
____

B. TEACHING STATISTICS

4. How long have you been teaching CSYS Statistics? ________ years ________ months

5. Number of students in your class sitting for CSYS Statistics (paper III) in
1992________ 1993 ________ 1994 ________

6. Recommended and additional textbooks used in teaching CSYS statistics. (please state names and authors of books)
____
____
____

For Question 7, please refer to the following table.

Please try to answer this question as much as possible. It is important for the researcher to know which of the following topics teachers find difficult to teach and also the topics that students find difficult to understand.

7. Please indicate teachers' level of difficulties in teaching the topics by placing a (✓). Please indicate students' level of difficulties in understanding the topics by placing a (X).

Topics/Concepts	Very Difficult	Difficult	Neutral	Easy	Fairly Easy
1. Populations and Sampling					
a. Definitions of populations					
b. Random sampling - simple, stratified, cluster					
c. Design of simple surveys					
d. Dot plots, Box plots					
e. Stem and leaf diagrams					

f. Mean, median and mode					
g. Interquartile range, variance and standard deviation					
h. Cumulative frequency					
2. Bivariate Samples					
a. Scatter diagrams					
b. Regression line					
c. Correlation					
3. Probability					
a. Random events and theoretical probability					
b. Sample spaces					
c. Equally likely outcomes					
d. Mutually Exclusive Events, Complimentary Law					
e. Addition Law, Multiplication Law					
f. Conditional Probability, Independence					
4. Probability Distributions					
a. Random variables					
b. Mean and variance of the sum and difference of independent random variables					
c. Binomial distribution					
d. Poisson distribution					
e.Poisson approximation to Binomial distribution					
f. Normal distribution					
g. Normal approximation to binomial, continuity correction					
5. Confidence Intervals and Hypothesis Testing					
a. Confidence Limits					
b. Simple ideas of significance tests					
c. Null and alternative hypothesis					
d. Critical region and significance level					
e. Sign Test					
f. Mann-Whitney (Wilcoxon Rank Sum) Test					
g. χ^2 - test					

8. Are there any other difficulties that you faced in teaching the subject?

Yes ________

No ________

If Yes, please explain briefly.

C. TEACHING METHODS

9. (a) Do you use the following methods in teaching CSYS Statistics?
Please tick (✓) as many as necessary.

Teaching Methods	Use it often	Use it sometimes	W ill likely be using it in the future	Won't likely be using it in the future	Will not use it	9.(b) ✓
1. Lecture						
2. Investigative approach						
3. Individual project						
4. Group project						
5. Classroom practicals						
6. Tutorials						
7. Computer exercises						
8. Problem solving through * small group discussion						
9.** Large group discussion						
10. Others (Please specify)						

* small group (less than 10 pupils)
** large group (at least 10 pupils)

9. (b) In your experience, which of the above methods do you think could help students to understand basic statistical concepts?
[place a tick (✓) in the last column of the table in 9.(a)]

D. TEACHING MATERIALS

10.(a) Do you use the following materials in teaching CSYS Statistics?
Please tick (✓) as many as necessary.

Teaching Materials	Use it often	Use it sometimes	Will likely be using it in the future	Won't likely be using it	Will not use it	10(b) ✓
1. Recommended textbooks						
2. Own teaching notes						
3. Teaching materials developed for school teachers						
4. Problem sheets						
5. Test Bank of Questions						
6.Support materials from Scottish Examination Board (SEB)						
7. CSYS past examination questions						
8. Published practical teaching materials						
9. Published journals and articles						
10. Materials gathered from courses and workshops						
11. Case Studies						
12. Others (please specify)						

10.(b) Which of the above teaching materials do you think could be used effectively to reinforce basic statistical concepts.
[place a tick (✓) in the last column of the table in 10.(a)]

11.(a) Which of the following methods of assessment do you use to measure students' understanding of the subject? Tick (✓) as many as necessary.

Measurement	Use it often	Use it sometimes	Will likely be using it in the future	Won't likely be using it	Will not use it
1. Topic test					
2. Test after each lesson					
3. Talk to individual students					
4. End of term exams					
5. CSYS exams					
6. Projects					
7. Classroom practicals					

11.(b) Please specify other assessment methods that you have used to measure students' understanding of the subject.

____

12. Why do you think students find certain basic statistical concepts difficult to understand?
Give one concept (other than Probability) as an example.
(Use a separate sheet if necessary)

____

____

____

____

13. Please provide any comments and suggestions about the study and the questionnaire.

____

____

____

14. Please delete one.

I **may / may not** be contacted in the future.

Contact address: ____________________

____

Telephone no: ____________________

Thank you for your cooperation.

Appendix A7(ii): Questionnaire to Lecturers of Statistics

Lecturer of Statistics
Department of Statistics
University of Glasgow
Glasgow, G12 8QQ

Dear Sir/Madam:

A Study on Teaching Statistical Concepts at Introductory Level

I am currently engaged in a study on methods of teaching statistical concepts at the introductory level. Part of the study is to look at the effectiveness of various teaching methods used in teaching statistics at several colleges of higher education and universities in Scotland; the difficulties in teaching certain concepts; why students find certain concepts difficult to understand; factors that contribute to the teaching and learning difficulties of the subject and many other enquiries which led to the teaching and learning of statistics.

The main purpose of the study is to find out whether the use of various teaching methods in teaching statistics can contribute to students' understanding of some difficult statistical concepts and that certain methods of teaching might help to diagnose the problems that students are facing in learning the subject.

My aim is to develop a statistical teaching model that could be of use in the teaching of certain concepts effectively through several experiments which have been conducted recently at the University of Strathclyde. However, much investigation and co-operation are needed from college and university lecturers before any firm conclusions can be made about the model. I strongly feel that statistics lecturers would provide a better insights into the teaching of statistics and that their contribution could prove to be beneficial.

Enclosed please find a questionnaire which I hope you could fill in and return in the prepaid envelope as soon as possible. I will assure you that all information given by those who are directly or indirectly involved in the study will be treated with utmost confidentiality. I wish to convey my thanks beforehand and would welcome any comments and suggestions regarding the study and the questionnaire.

Yours sincerely,

Zamalia Mahmud
PhD Research Student

Please answer the following questions to the best of your knowledge. All information given in the questionnaire will be treated with utmost confidentiality.

A. GENERAL

1. Name of lecturer: (Ms/Mr/Mrs/Dr/Prof.) ____________________
2. Name of College/University: ____________________
   Department: ____________________
3. Highest degree of qualification (specify area of study): ____________________
   Other qualifications: ____________________

B. TEACHING STATISTICS

4. How long have you been involved in teaching statistics at introductory level?
   __________ years/months
5. Recommended and additional textbooks used in teaching introductory statistics. (please state names and authors of books)
   ____
   ____
   ____

*For **Question 6**, please refer to the following table.*

Please try to answer this question as much as possible. It is important for the researcher to know which of the following topics teachers find difficult to teach and also the topics that students find difficult to understand.

Mark on the topics which is applicable to your syllabus or teaching module only.

6. Please indicate teachers' level of difficulties in teaching the topics by placing a (✓).
   Please indicate students' level of difficulties in understanding the topics by placing a (X).

Topics/Concepts	Very Difficult	Difficult	Neutral	Easy	Fairly Easy
1. Populations and Sampling					
a. Definitions of populations					
b. Random sampling - simple, stratified, cluster					
c. Design of simple surveys					
d. Dot plots, Box plots					
e. Stem and leaf diagrams					

f. Mean, median and mode					
g. Interquartile range, variance and standard deviation					
h. Cumulative frequency					
2. Bivariate Samples					
a. Scatter diagrams					
b. Regression line					
c. Correlation					
3. Probability					
a. Random events and theoretical probability					
b. Sample spaces					
c. Equally likely outcomes					
d. Mutually Exclusive Events, Complimentary Law					
e. Addition Law, Multiplication Law					
f. Conditional Probability, Independence					
4. Probability Distributions					
a. Random variables					
b. Mean and variance of the sum and difference of independent random variables					
c. Binomial distribution					
d. Poisson distribution					
e. Poisson approximation to binomial distribution					
f. Normal distribution					
g. Normal approximation to binomial, continuity correction					
h. Sampling Distribution of the Means					
5. Confidence Intervals and Hypothesis Testing					
a. Confidence Limits for population means and proportions					
b. Significance testing: Z-test					
c. Significance testing: t-test					
d. Null and alternative hypothesis					
e. Critical region and significance level					
f. Sign Test					
g. Mann-Whitney (Wilcoxon Rank Sum) Test					
h. χ^2 - test					

C. TEACHING METHODS

7. (a) Do you use the following methods in teaching introductory statistics? Place a tick (✓) where applicable.

Teaching Methods	Use it often	Use it sometimes	Do not use at all	7.(b) ✓
1. Lecture				
2. Investigative approach				
3. Individual project				
4. Group project				
5. Classroom practicals				
6. Tutorials				
7. Computer exercises				
8. Problem solving through * small group discussion				
9.** Large group discussion				
10. Others (Please specify)				

* small group (less than 10 pupils)
** large group (at least 10 pupils)

7. (b) From your experience, which of the above methods do you think could effectively help students to understand basic statistical concepts?
[place a tick (✓) in the last column of the table in 7.(a)]

D. TEACHING MATERIALS

8.(a) Do you use the following materials in teaching introductory statistics? Place a tick (✓) where applicable.

Teaching Materials	Use it often	Use it sometimes	Do not use at all	Not available	8. (b) ✓
1. Recommended textbooks					
2. Own teaching notes					
3. Problem sheets					
4. Teaching materials developed for college/university lecturers					
5. Test bank of questions					
6. Past examination questions					
7. Classroom practical teaching materials					
8. Published journals and articles					
9. Materials gathered from courses and workshops					
11. Case Studies					
12. Others (please specify)					

8.(b) Which of the above teaching materials do you think could be used effectively to reinforce basic statistical concepts?
[place a tick (✓) in the last column of the table in 8.(a)]

9. Please provide any comments and suggestions about the study and the questionnaire.

____

____

____

Thank you for your co-operation

Appendix A8(i): Open-ended Interview Questions

Note to reader

These interview questions were constructed in accordance with the research objectives and aims of the interview. However, in certain circumstances, some questions were not directly related to the research objectives. These questions were intended to probe some of the vague responses and to help the interviewer in obtaining specific information from the subjects. Some of these questions are not stated here.

A. The difficulties in teaching certain concepts

- Why do you feel that 'this concept' is a difficult (or easy) concept?
- Do you find 'this concept' difficult to teach?
- Briefly explain how do you go about introducing the concept for the first time?

B. The aspects of teaching methods used and the purpose of using them

- You use 'these teaching methods' quite often. Why do you use them?
- Do you find that they are useful? Why?
- Explain briefly how you use 'these teaching methods'?
- How can these methods be used to help students to understand the concepts?
- You also use 'these teaching materials' quite often. In what circumstances do you use them?
- How can these materials be used to reinforce students' understanding of the concepts?

C. The effectiveness of teaching methods used

- Do you think the methods which you often used in teaching can be used effectively to reinforce students' understanding of the concepts?
- To what extent do you think these methods have helped your students in understanding the concepts?

D. Methods of evaluating pupils and students' understanding and misunderstanding of the concepts

- Generally, how do you evaluate pupils' and students' understanding and misunderstanding of the concepts?
- How do you know if your students have understood (or not) the concepts? Mainly, what is the signpost that they have understood (or not) the concepts?

E. Factors that contribute to students' difficulties in understanding the concepts?

- What are the factors that you think have contributed to students' difficulties in understanding the concepts?
- Why do you feel that 'these factors' can contribute to students' difficulties in understanding the concepts?

F. Factors that contribute to the subjects' difficulties in teaching the concepts?

- What are the factors that you think have contributed to your difficulties in teaching the concepts?
- Why do you feel that 'these factors' have contributed to your difficulties in teaching the concepts?

Appendix A8(ii): A Sample of an Interview Transcript

Name: Lecturer 7
Qualification: MSc (Statistics)

Q1. Which of the concepts do you think students find difficult to understand?
A1. I think students find anything to do with probability very difficult. They can start off feeling very unsure of what statistics is. I think they find hypothesis testing more difficult than say idea of the point estimation. And within hypothesis testing I think they find the non-parametric methods quite difficult to do and to understand.

I teach things like sampling and experimental design at an elementary level. I've always enjoyed doing it. Students who are more mathematically inclined find these topics particularly difficult often because they don't have any practical experience to back up what we're telling them. So, if you're teaching sampling to people who are doing economics or sociology, they have an understanding of why you're into these kinds of things. If you're teaching experimental design to Biologist, they can see the relevance of it. But if you're teaching these ideas to mathematics students, then they really don't have this practical background to understand why they're doing it or why they might want to be involved in increasing precision for the same sample size or increasing power.

Q2. What are the factors that contribute to students' difficulties in understanding the concepts?
A2. They tend to find the approach of hypothesis testing quite an alien one for reasons that I'm not altogether sure about because if you like we would think about scientific method as being similar to what we called hypothesis testing that you start any scientific study with a question. So, I've never been too sure but they do find that very difficult.

We tend to teach in terms of statistical inference. We would do exploratory data analysis with maybe a bit of probability, just ideas of what probability might be. We would then talked about point and interval estimation if we were doing a hypothesis testing. Now we didn't used to do that. We used to do hypothesis testing and then interval estimation and it's still done that way in quite a lot of places but we find that the idea of an interval estimate is a very natural one and students find that very straight forward but there seems to be a difficulty going from the idea of the interval estimation to the idea of hypothesis testing. I've never been absolutely sure why that should be.

I would suspect that part of it is because we try to introduce an awful lot of new concepts all at once because its a whole idea of null and alternative hypothesis. We might be bringing in significance levels and p-values and as I say they could find probability quite difficult, so immediately you introduce some probabilistic concepts there. There's quite a lot coming in all at once. They sometimes muddy the waters with one or two sign test and that really caused the problems at the beginning. And unless you think very carefully about how you're introducing these ideas, then you can be in difficulty.

One thing that we find has helped is having some idea about point estimation and fairly well developed idea of interval estimation. In outcome to hypothesis testing, say, well, we can go to these problem of interval estimation in a slightly different way. We can think of it in a hypothesis testing framework to see evidence support these ideas that we have that there is no difference between these two populations. Then there's a kind of a heuristic buildup of the idea of null and alternative hypothesis. So you can kind of build up a bit from the interval estimation side.

So I don't know really why students have this difficulty in hypothesis testing. I think partly is because they comes to early on our syllabus before they've done simple ideas of interval estimation, sampling variation and I think partly because so many concepts were brought in at once. This muddying of the waters with one and two-sided test and you have to think particularly through the presentation of it. There's a sense of throwing things a bit. There's a whole lot of ideas to give out in this one whole lecture.

You know something out of the blue, you haven't done anything but exploratory data analysis, probability .., something that come as this whole framework of null hypothesis, of alternative hypothesis, sig. levels, why rejecting Ho when you should know its importance, why we concentrate on one error rate and not the other. It's always a mystery to students. So, I think you have to take longer to explain hypothesis testing.

Q3. Do you have the problem of finishing the syllabus, rushing through things...
A3. Not so much now. We've been changing our courses quite a lot in the last few year's. We've divided students into different classes depending on their mathematical ability. We rewritten the syllabus a number of times trying to take into account what they should know in school, about their background. So we are trying to direct what we teach, so these has cut down a bit on the problems of getting through a wide general syllabus, say we are not going to bother the biologist with time series, the social scientist with repeated measures. So I think there is not such a problem in that.

Q4. From your experience, do you yourself find certain concepts difficult to teach or trying to get students to understand certain concepts?
A4. Yes, I spent a lot of time getting a lot of basic concepts much more than I would at one time. The basic concept, i.e. sampling variability is something that I would have at one time thought of as pretty obvious. If you got a fixed population, if you take different samples from them than you have certainly different characteristics. That I find even with students who have an experience of a practical subject like Biology, that concept is not always immediate to them. So I spend a lot of time talking about repeated sampling, simulated repeated sampling, looking at the effect of the samples of size on spread of the answers that you get, leading to the discussion of a standard error, discussion of an interval estimation.

I find that students respond better to that kind of approach or simulation, step by step through the whole questions of sampling variability and into interval estimation from there. Hypothesis testing I would have admit I am never particularly happy with myself because I think there's so many philosophical problems, then there's the classical statement of hypothesis testing you ignore one error rate now trying to explain to students that's what you really wanted to do. It's really difficult because it's no longer clear to me that's what you really want to do, so I find that a bit trivial.

They find problems of paired data difficult. The questions of when are the samples independent and when are they paired. Now again I'm never too sure why they find that difficult. It's not a difficult concept. The paired because you got two measurements for the same individual. Why does that make a difference? It's so strange that students not being able to spot that kind of problem. So again it take quite a lot of time to do examples of paired and independent data.

The regression and correlation, now conceptually students don't have a great of difficulty. Again they have difficulty knowing which is which.

Q5. Are the students happy with the methods that you use in teaching them?
A5. Mainly we lecture just in a traditional way. I've been trying over the last couple of years to vary that quite a lot and the course design now involves students with the less lecturing and more practical work. So they're learning Minitab fairly intensively in the first year. So, that's help. Instead of two lectures a week plus occasional labs. They've now got three lectures a week plus a weekly two hour lab - that's the first year course. So that makes a difference to how they're learning. You no longer teaching them facts. I think we're trying to impart skills as well. I think that's so much better way of teaching statistics but of course that wasn't an option that was open to us until recently because the basic conversant depend on hardware and software. So it's a big investment at the time but it's worth it at the end because students really appreciate doing something seeing how it works out in new context rather than passively receiving the information in the lecture.

And also I've used videos in my lecture last year. Split them up into little groups, discussed things doing lecture. I've done occasional experiments, I've got them to do things during the lecture and discussed the results so that notes are printed there for them while when we're doing the lecture isn't necessarily simply bring the notes together.

I wish I could say that has being better received by the students in the past. Now I was a bit nervous about it all, I've never tried to split them into discussion groups. They were doing well in the class but by the time we ask their opinion about the whole thing at the end of the course, they seem to be a bit kind of uncertain. But that was a lot of factor involved in that, of course. I felt that the students in the end understood the concepts better as a result but it would be difficult to measure either by the response or by the results in the exams.

It's an instinct of looking around and of hearing. Maybe you don't often get a chance to hear students' response to things in a traditional lecture. You don't hear much from the students. In the way I was doing things last year, I was hearing what the students had to say. We had a discussion and then we were trying to have individual discussion regarding notes from the board or whatever. So they had to speak back to me to hear from me. So I had a pretty good impression from that. They really don't understand what was going on.

Q6. Do you have any kind of formal evaluation?
A6. Yes, they did project last year. Again, for the first time we were dealing with continuous assessment and they had to hand in reports and they had to do a couple of projects. I must say the projects were surprisingly good and understood some fairly good concepts. The projects was on time series. I was really very pleasantly surprised by the quality students' attempts and they also did a lot of work on lab report. And during the practical session, a vast majority of the students seem to have a pretty good idea of what they are trying to do each time.

But I would say exam performance was weak. So, I was quite pleased with what they do in their projects and I feel that the students' reaction was that this was all a bit alien... why are we doing these exercises, why can't you just lecture to us. Students are very conservative in my experience. They are very much more conservative than the members of staffs. They expect things to be done in a particular way and if you deviate from that they get very annoyed.

Q7. Were they not told about why they're doing these and that?
A7. They were told about it. It was more practical and the focus of the course is being able to do things as well as understanding some of the things.

Q8. When you get them to do the project, do they do it in groups or on their own?
A8. Well they each hand in an individual report. I would suspect in the labs that they naturally work in twos or threes, well that seems to be fair enough.

Q9. Do you think it's good for them to work together?
A9. I think it's good for them. Of course there's a considerable difficulty in the assessment if don't hand in an individual project report.

The honours students work in groups to do projects. These projects do not count for assessment. They do work on some problems in groups and they enjoy that and they enjoy the experience of working together.

Q10. How many students are you roughly talking about in one class?
A10. In first year, this year just about 100 in one course. Next year that number will be up to maybe 250 because of various structural changes that are taking place. So we are going to have a bigger problem. I think we were glad to have only a hundred students first year through out a particular course. And it looks so it is going to be a lot next year.

And so the logistics of organising that small groups labs every week for all these students are going to be a difficulty.

Q11. How do you vary your teaching methods among the concepts? In what circumstances would you vary your teaching methods?
A11. For probability I find that... well, I wouldn't hand out notes for probability. I would actually lecture. We do things on the board with plenty of examples. Because I think the students would think it very difficult and the problem with printed notes and handouts is you get through more material in the same amount of time. They also find probability a bit dry and boring. So if you hand out notes, they get bored. So you best just going through with them. So everybody was going along at a certain speed and they're actually seeing examples and see some theory buildup. With the statistics on the other hand, plenty of handout notes and doing some examples.

Q12. What about concepts on inferential statistics - confidence interval?
A12. I use quite a lot of simulations. One of the problems is making those happen live in a lecture. That is something that I have to get sorted out this year.

For the more advanced group I think it's appropriate to work at a theoretical foundation and expect them to be able to cope with that and to see the implications rather than with the elementary groups where simulations can help you to avoid a theoretical derivation of things. For example, you don't really need to worry about central limit theorem or whatever for what happens to the mean for a set of observations. Simply increase the number of observations in your simulation.

Q13. Of all the methods that you use, which one favoured you the most? Which one you think would help them to understand better?
A13. Getting the students to do anything themselves by far the best way to get across ideas for example in sampling. I enjoy teaching sampling design but as I said to you earlier I think some students can find it difficult to see the point. Why is this probability sampling, why does it matter?

A little exercise I came across this year and I've used with a lot of groups, undergrads, postgrads, ... is you get a page and you get 60 circles on it and the circles are of different sizes. I said to students, each of the circle has got a diameter which is divisible by 0.5 cm, pick 5 of them and find the mean. So they all choose a sample of 5 circles and they find the mean. What I found was, this is the mean of the sample, they were getting answers any place between just over 1 cm and about 2 cm. Not a single person got an answer below 1 cm in a class of almost 80. That was at the beginning of the lecture. At the end of the lecture, we develop the idea of probability sampling, random digits and simple random sampling. So then you got them to get a random sample of 5 circles by numbering them using random number tables and we got a different results.

This time we find the answers they got were very evenly divided around the value of 1 cm. The reason being that the true mean of the circles on the page is 1 cm. So the next lecture started with me putting up a graph (because I collected the results) of the two experiments and of course everybody is sitting there with this bias results on the first experiment and this nice distribution around the true value on the second.

As a point I say, well you see, convenience sampling is very dangerous. You can end up with very bias results and that's what probability sampling is trying to do, to avoid having bias results. It really make a very big impact every time I did the experiment.

I did it with school teachers who are teaching statistics and they get the same results as the students and when I pointed this out, you can just see that something... a very powerful, very simple idea but getting the students to do themselves, to do the sampling experiment. Very simple one it makes more impact than just showing them the results because they are actually doing it. So they realised afterwards what mistakes they've made.

Q14. What is the most difficult aspect of teaching statistics at this level?
A14. The thing I find most difficult in elementary statistics is non-parametric test. You have a much better idea in what teachers find most difficult than me. But that is something I find particularly difficult.

I become very aware of that recently and the reason is because we have taken non-parametric testing out of our first year syllabus 3 years ago. In that space of time, I become much more involved with school teachers and the sixth years school syllabus in Scotland is basically non-parametric and perhaps that had made me realised that I'm very wary of this non-parametric testing. I think it can be a very difficult idea for students.
I know I find it very easy to motivate students towards t-distribution or at least the idea of point and interval estimation in a non-parametric sense. A mean is something that is fairly natural to students to look out but you do have to warn them about occasions when it's not really representative of... So you're going direct with mean as a point estimate to looking at an interval around about the sample mean as .... and then going into t-interval and t-test.

It seems to me that the equivalent non-parametric procedures aren't that straight forward to motivate because you've got this rank based approaches. I think the rank based approaches are conceptually very difficult. They can see the idea of ordering it, that makes some sense but as soon as you start to order, your confidence interval is being build up in a matrix equation depends on the ranks, so you have actually abstracted away from the data itself to the ranks of the data. You're building up your confidence interval depending on the rank and then you have to convert back to look at what is that rank, what is three and a half observation means. I think students find that one quite difficult.

With non-mathematical students the whole theoretical basis can be very difficult but the thing that really worries me about the non-parametric approaches at an elementary level is with students end up with idea that these non-parametric approaches are always valid which isn't true.

if you think about Wilcoxon sign rank test, you're assuming symmetric. Now that message is a lost and if you look at many textbooks, the reason for it being lost is quite simple...they show some skewed distribution. Well, say, this is a sample. This is obviously non-parametric, therefore we can't use t-interval. Well, we just go for this Wilcoxon sign rank test and that's not right. If you think of the Mann-Whitney test, the assumption there is of equal variances and again that's a message that doesn't get through to the students.

So I always end up with this feeling for the students especially if you teach both the parametric and the non-parametric. The parametric is quite straight forward that is for normally distributed data or the central limit theorem.

I have this debate with people involved in the sixth year school syllabus which is why aren't you doing t-test and there doesn't seem to be a good answer except that the syllabus was created in that period in the 80's when there was a tremendous enthusiasm for non-paramertic test. But there was a kind of huge...for non-parametric analysis and this sixth year syllabus was created in that time and I'm not convinced why.

The other problem is when they're talking in schools about developing further statistics, where do you go? The best place to go is into regression and then to confidence interval and hypothesis testing related to regression problem. Well they do regression but the only thing that they do is fit a line.

So I was telling them why not teach them to test whether the regression line is a good fit whether parameter is different from zero. If you develop t-test before hand, that's straight forward.

Q15. Do you think that one of the factor that caused the difficulty in students' not understanding the concepts is because they're not exposed to statistics in school?
A15. Yes. Scotland has got the least well-developed school syllabus for statistical concepts than any other country that I'm aware of. It's much poorer than England.
As you know, I'm getting quite involved in the schools and I think it could be seen something as a crusade. The Scottish schools has got this 5-14 syllabus which develops very nicely, the information handling strand of mathematics skill and there's also 5-14 environmental studies strands which targets vaguely statistical data handling issues of knowing what is an appropriate sample... So these develops really beautifully along second year secondary level at the 14 years old and then it stops, that's it and it's nothing until 6th year if they choose to come back and do it.

And that gap in the main stream standard grade and higher maths syllabus is a huge problem because it creates tremendous difficulty because in Biology, Geography, to some extent Sciences and Business Studies do this sort of things in standard grade and higher.

So it's extremely inefficient to have this individual subjects trying to deal the same basic data handling problems. Again when you get to 6th year level, you're then dealing with ideas that are called statistics in Paper 1 syllabus that are really, here's how you draw a boxplot or a stem and leaf plot,... here's what a median is..

And really you're talking about the 17 year olds, they think that statistics is tremendously boring and simple subject because they're only being taught at 17 and the things that are quite capable of knowing when they're 12 or 13. And the first time they come across probability syllabus could be at 17 because in Scottish school syllabus there are no probability in the 5-14 information handling and certainly not in the standard grade maths.

So you could have a generation of kids who first come across the concepts of probability in the age of 17...... One of the things we did this year was we produced standard grade maths investigations which deals simulation and simple probability concepts as a way of plugging the gap a little bit.

Q16. Is statistics not properly taught at school level?
A16. The school inspectors wrote a report on mathematics teaching in which the area of criticism is on data handling and they compare with kids in the other countries that school children in Scotland do not have the right exposure to data handling and data collection.

They wrote a report a couple of years after the 5-14 syllabus in which they actually gave a summary of their rating of schools primary and secondary on all the aspects of the 5-14 syllabus and whereas all primary schools did reasonably well in all the areas of the 5-14 syllabus. Secondary school fell down very badly in the 5-14 information handling. And they commented that the major problem they saw was a lack of progression in data handling in secondary school.

I spòke to one of the HM inspector here, a man with a particular interest in Statistics and his comment was he's still seeing the same graphs in second year secondary classes as he have seen in Primary 5 and I would say that all you can assume that school children in Scotland learned about statistics....

That means that our first year university courses, we have to start from there. So we have to teach them about boxplots, about stem and leaf diagrams, about medians and about probabilites. And it is not impossible that a large proportion of people sitting in their first year statistics course might never have heard about probability.
So that means we are starting at a very low level when you compare with the maths class or the geography class and that causes problems for us as well, the recruitment into honours statistics where we allowed people into 2nd year statistics who haven't done first year statistics as long as they've done first year maths at a university. And the reason is simply that a lot of people out there who come to the university to do the mathematical sciences have never heard of statistics, avoided in the first year and here we are quite anxious to encourage them to think about it in 2nd year.

So we still teaching boxplot in second year statistics, obviously we don't spend a lot of time on it but you still got to get this kinds of basic ideas across because you can't

assume anything from school. So there's no common language of statistics and I find that worrying because that means their vocabulary of statistics isn't being developed in school.
Teachers themselves having great difficulties with teaching statistics. I do sympathise with them. The problems are institutionalised not problem necessarily of individual teacher. This kind of lack of progression which the inspectors are complaining about and I said this to the teachers as well. In the sense who are they blaming for the lack of progression. It can't be the teachers necessarily in the classroom because if you look at the guidelines there is nothing to progress into because the people who are the 3rd, the 4th and 5th year syllabus of schools have not built in any progression.

You have heard how the Scottish school syllabus in maths has come to be developed. About 10 years ago, the Scot O-Grade which is equivalent to England O-level and they introduced the Standard Grade, so that came in 3rd and 4th year, that fixed in 3rd and 4th year school syllabus. They then add in the 5-14 syllabus in at the bottom end but certainly in statistical area it doesn't link in and they now try to sort out the 5th and 6th year.

Then certainly in the 5th and 6th year they discovered there're quite a lot of problems in the 3rd and 4th year level but they are not allowed to change those because those are fixed. So they now bolt in something from the top end into 5th and 6th year but the trouble is the 5-14 syllabus has been bolt in from the bottom end doesn't lead on to properly and the 5th and 6th year has been bolt in here, they can't go back to change the 3rd and 4th year.

So it's exactly the way not to create a unifying progressive syllabus and because statistics is seen by mathematicians as a 'frame' subject, statistics and statistical ideas has been particularly badly served by us. They've just fallen through the cracks.

So teachers are having to try to cope with the syllabus which is obviously inadequate. They don't have a statistical background. They've used maths teachers, sciences... So you've got a very largely unhappy teaching posts. They're aware of their own inadequacies and so they avoid it or they tried to teach statistics as though it was mathematics.

One of the things I was trying to get across the teachers this year in the workshop, there's a very basic difference between the two subjects because mathematics is deductive and statistics is inductive or inferential and you can't teach them in the same way, they just doesn't work. If you try to teach statistics as a deductive science then you've lost an understanding of what it really is.

'R' said that the maths teachers are used to having answers at the back of the book and he said that teaching statistics was like embarking on a journey where the destination is unknown. That was quite a nice idea of saying, well, we are going to look at something, we're doing something practical, we don't know what these people are going to find. We don't know what questions are going to spark off, what sort of comments we're going to make at the end of the day. So we don't have the score of the answers at the back of the book.

We've seen about 60 different teachers last year, I hope we have tried to encourage all of them to go away and think differently about how they teach statistical ideas whether it's 5-14, middle years or whether it's the 6th year. We tried to encourage them to do experiments to get them to do things practically, to think about issues of how you collect data and what kind of samples you're going to get, why it might not ..... , or how you're going to design the experiment. So the whole question of how you collect data and how that affects, how you analyse it. And I think people are trying now.

'R' and I don't particularly like the Scotvec modules in statistics. It's not the particular statistics one but it's the whole approach of Scotvec that we don't like because it's not really very educational, it's a kind of training of very specific skills.

I think the most depressing of all the experience of last year was the workshop we held for the teachers in the 3rd and 4th year of the Scotvec modules of realising that we were just talking of an alien language to people present who would come to find an answer at the back of the book. They were people who were obviously harassed. They were dealing with classes who were probably not very well motivated who didn't like maths as statistics is part of maths. They felt under resourced and under time pressure and because of inadequacies of the approach education they weren't very motivated themselves.

This is one of the activities last year that didn't work. It's a key area because they can do modules instead of the standard grade maths and it's dominated by this Scotvec modules. I wouldn't be motivated to teach the Scotvec modules.

Statistics is difficult in schools because to do something practical you really need two members of staffs present. Primary schools do statistics quite well in my experience because they were in topic based learning and you can introduce ideas of graphing data, measuring things and so on. Fairly natural in that context and there's usually in Primary school. I'm aware now there are members of staff floating. There are learning support teachers or something who can be brought in if there's anything going on...to take group of kids. The primary schools I think are doing pretty well on the basic collecting data but they falls apart in secondary.

One of the stupidest ideas in my opinion is this idea that they start in secondary 1 into maths with unstreamed classes, so you can have people who are in the same group of 30 kids. You can have some people who have done write up to level D or E, the 5-14 syllabus, whereas you might have people who for one reason or another have barely done level C of the 5-14 syllabus. So you've got a tremendous variation in their experience, in their ability, interest and there will be one teacher who is going to try and sort the whole lot out. I just find that amazing.

Appendix A8(iii): Matrix Displays 1, 2 and 3; Summary Tables 8.5, 8.6, 8.7 and 8.8.

MATRIX DISPLAY 1 of Conceptual Framework 1

Level of Experience	Subject	Factors that contribute to pupils and students' difficulties in understanding the concepts	Factors that contribute to teachers and lecturers' difficulties in teaching the concepts	Areas of Misconception
10 years or less				
3 years	T2	**PKNOW**: "Pupils still think statistics is an extension of mathematics. Their main objective is to get a numerical answer rather than try to interpret it. It is different from everything else they have ever done before and this is their first time learning statistics in a conceptual manner."	**LEX:** Teacher have little experience with certain concepts like survey design and sampling.	**THEO**: Fail to formulate null and alternative hypothesis. Cannot differentiate between a one-tailed and two-tailed test. **COMP**: Arrived at a numerical answer too quickly-no basis of understanding. Prone to careless mistakes. **INT**: Could not interpret results of hypothesis testing
3 years	T3	**PKNOW**: "Pupils have not been exposed to any kind of statistics syllabus before."	**LCON2**: Teacher finds statistics more difficult than any other subjects. **LEX**: "In the first year. I was not sure how to go about introducing each concept.I find it very difficult to be sure of doing the right thing"	**THEO**: The notion of the degrees of freedom. the purpose of finding them. what exactly do they have to do with it. The notion of the null hypothesis. what to test. **INT:** What to do with the results
8 years	T4	**PKNOW**:"It doesn't relate to every thing else they have done in Maths" **ABLE**: Some pupils do not have the natural ability to assimilate ideas." **ATT**: Didn't do background reading themselves **TF1**:"They need more time to think about the concepts and problems related to the concepts."	**LEX**: "I don't fully understand everytime I do it and tend to accept formula from the book."	DK-3

DK-3: Ambiguous answer

Table continue on next page.

7 years	T5	**PKNOW**: The first statistical concept they meet. **ABLE**: Some pupils are not able to think through the ideas quickly. Some learn by rote. **TF1**: Need more time to come to terms with statistical procedures. Everything had to be rushed towards the end of the course. **ATT**: Some couldn't be bothered and would rather be told what to do	**LCON2**: No confidence in handling difficult concepts.	**THEO**: Fail to formulate the null and alternative hypothesis. **COMP**: "Pupils often rush into the calculation part. They are keen to end the calculation." Prone to making mistakes. **INT**: Could not interpret results properly.
2 years	T6	**PKNOW**: "They need a strong and good foundation lower down the school to allow statistics to grow on them, to be more familiar with the ideas later on. If they have gone through this before, it will help them and have some standing with them" **ABLE**: Some are not able to think through the ideas quickly. **TF1**: "They need more time to let the ideas grow on them. Given more time they may have a greater awareness and a greater understanding"	**RES**: "Materials on certain concept like Mann-Whitney test are very poorly resourced-makes it more difficult to understand and teach "	**THEO**: "They don't understand the meaning of significance level. They used in sentences because they were told to qualify their answers"
1 year	T8	**ABLE**: Pupils tend to learn by rote, thus no real understanding of the concepts.	**LEX**: Teacher is new at the subject. Not knowledgeable on subject matter. Find difficulty in tackling questions posed by pupils or in providing a satisfactory answer. **LCON2**: Fear of being incompetent and lacking in confidence, teacher tend to learn by rote as well.	DK

DK: Don't Know

Table continue on next page,

1 year	T10	**PKNOW**: "Pupils are not used to statistics, they have never done it before. They said everything is just so new. They don't understand why they are doing it" **ABLE:** Some pupils are not able to think through the problems on their own. They are poor at their own initiatives. They are too exam-oriented. **TF1**: Little time is invested in doing practicals and investigative approach that would help pupils to understand better. **ATT**: "Pupils are used to being spoonfed. They are used to having things on the board"	**LEX**: First time teaching the course. **PRESS**: "I don't enjoy teaching the later part of the course. Teaching are very disjointed .... course had to be rushed at the end....not happy with my approach in teaching the concept"	**THEO**: Find difficulty to set up the null and alternative hypothesis. deciding which test to use: one-tailed or two-tailed. Not sure what is a standard deviation. what does it measure. **COMP:** Weaker pupils have difficulty in understanding how to do simple mathematical calculation like factorials. algebra. straight line equation. **INT**: Poor at interpreting results and interpreting graphs. "They are not good at explaining why. .....not happy putting English skills into Maths paper"
2 years	T11	**PKNOW:** Lack of mathematical background. **ABLE**: Not capable of understanding what was taught. **TF1**: Due to time constraint, pupils are not able to do practicals and investigative approach that would help them to understand better.	**LEX**: First time teaching the course. **LCON2**: Not confident enough since teacher is at the stage of learning and teaching himself as well as the pupils. **TF2**: Lack of time to look for more or new resource of information. Need more time to prepare teaching materials. **RES:** Resources not easily available	**THEO**: Find difficulty to set up the null and alternative hypothesis, deciding which test to use- one-tailed or two-tailed. In sampling, the difficulty is in choosing the appropriate sampling methods. **COMP**: Difficulty to grasp the mathematical notations
3 years	T12	**PKNOW**: "When they started to do statistics, they wanted to be very black and white. They wanted to be certain of everthing. They relate it to Maths , used to it as being certain" **ABLE:** Pupils wanted to be taken through step by step....cannot figure things out by themselves.	**LEX:** Difficult to teach in the first place, forgot everything that was being taught at the university - had to relearn everything. **LCON2:** Statistical ideas are quite complicated and difficult to get over them-need more confidence to tackle statistics questions. **PRESS**: Course had to be rushed towards the end to get through the work for the exam.	**THEO**: Difficulty in understanding critical region and relating to hypothesis. The idea of continuity correction...they forget more than they find it difficult. **COMP**: Rush into calculation, they are happy with just the numerical answer. No basis of understanding. **INT**: Do not bother to interpret results.

Table continue on next page,

2 years	T14	**PKNOW**: Pupils are not introduced to statistics earlier on. No emphasis of statistics in the 3rd and 4th year of secondary school. **ABLE**: Some are less able to think through the problems on their own, need more coaching. Most are mathematically less able, some simply don't like Maths. **TF1**: Shortage of time. Given only two hours per week-could not do practicals and investigative approach in which are thought to be among the best ways to get pupils to understand concepts.	**LEX**: Teacher find difficulty to put ideas across to pupils. Teacher felt approach to teach certain concepts is not appropriate.	**THEO**: Not knowing which test to use, one-tailed or two-tailed. Unsure of statistical tables to use, the levels of significance, critical region. They don't know where it is coming from and whether they use the right procedure to test.
10 years	CL5	**PKNOW**: "Science technicians who did statistics as part of the course found it to be very difficult because of no statistical background." 90% of students in this college have no statistical background and were not taught at school. **ABLE**: Students are not the best of class. No math background, numerically they are weak.	**TF2**: "Statistics is a nice subject if you can hurdle the time. If you can increase the time, then you can do experiments and practicals to make it a bit more lively." **PRESS**: No pressure. "Statistics has always been the second cousin and it is always been something that we will do it if we can and we won't worry about it if we can't"	**COMP**: They find counting difficult to cope. Numerically they are weak.
3 years	CL13	**ABLE**: Students cannot cope when too many concepts are thrown at them at one time. **TF1**: They need more time to sink in. There's too much in the syllabus to be covered within certain period.	**PRESS**: Have to follow standard of module. Tend to teach along those lines. There's not much leeway for experimenting and trying new ideas. **TF2**: Time are enhanced with modules. Need more time to teach difficult concepts.	**COMP**: "The students didn't understand how the calculations are performed, basically they don't understand what we are doing"

Table continue on next page,

8 years	L10	**PKNOW**: Too mathematically inclined- they are not used to any practical experience to back up what they are doing, to understand why they are doing it. They were not exposed to statistics in school. **ABLE**: They find concepts too difficult to grasp. Difficulty of concepts do not match with their ability to cope. Students are also introduced an awful lot of new concepts all at once-unable to cope. **TF1**: Students need more time to understand. **TCF**: Statistics were not properly taught at school-teacher don't have enough statistical knowledge and experience to teach the subject.	DK-1	**THEO**: "Students find the idea of an interval estimate very straight forward but there seems to be a difficulty going from the idea of the interval estimation to the idea of hypothesis testing. Paired data- the question of when the samples are independent and when they are paired. Regression and Correlation-knowing which is which.
11-20 years				
20 years	T1	**PKNOW**: Statistics were not taught in Standard Grade and barely in S5. No progression in the subject. Pupil first learn about probability theory in S6. **ABLE**: Unable to relate from one concept to another. Don't remember what it was to start with. Just accepting what is being said. **ATT**: Laziness. Simply don't like the concept to start with, for e.g Probability Theory.	**LEX**: Teachers' experience with simple surveys are so little. Tends to teach little or not teach at all.	**THEO**: Cannot decide on the meaning of the null and alternative hypothesis. **INT**: Pupils often do not relate back to the hypothesis when they are making conclusions about the test. Their conclusion tends to be very vague and that's because they don't actually remember what it was to start with.
17 years	T9	**ATT**: Pupils don't do much background reading, not conscientious enough. They expect differently of the subject-start to lose interest in it. **TF1**: Shortage of time, syllabus had to be pushed which results in pupils not being able to cope.	**TF2**: Shortage of time, syllabus had to be pushed on resulting in teachers not able to explain concepts in detail. **PRESS**: Syllabus is tight - pressure in completing the syllabus.	**COMP**: Weak in handling numerical problems. Prone to making careless mistakes. **INT**: They want to be able to do the process without necessarily understanding the concepts behind them.

DK-1: Question not asked of respondent

Table continue on next page.

12 years	T7	**PKNOW**: They have very little statistical background. If statistical ideas were introduced at an earlier stage, pupils may find it easier to get into grips with the ideas. **ABLE**: The weaker pupils have difficulties to think through the ideas without assistance from the teacher. **LCON1**: Pupils not trained to be more competent before they reach 6th year studies. They need more support and encouragement from staffs in the other departments	DK-1	DK-2
5 years	L2	**PKNOW**: Students do not have a strong mathematical background. **ABLE**: Unable to relate from one concept to another. Keep forgetting what was taught in the previous lectures.	DK-1	**THEO**: Difficulties in understandg Type 1 and Type 2 errors- unable to relate them to the hypothesis, and the decision rule. **COMP**: Poor in dealing with omputational exercises. Prone to mistakes.
20 years	CL6	**PKNOW**: They have probably never come across any statistical concepts before. **ABLE**: The number of concepts that they have to take in at one time and all the different formulas -tend to confused them. Some just don't have the ability to think mathematically. **ATT**: They don't spent enough time trying to familiarise themselves with the concepts and formulas. They finish up by rushing things at the end. **TF1**: "I think it's the case of time. If they do it over a long period of time, they eventually get the problem sorted out."	**TF2**: Teacher had to spend more time going over basic arithmetic which affects the time to teach more concepts.	**THEO**: Why the sample mean is just an estimate of the population mean. Unable to distinguish between different types of dists., which one to choose in a particular circumstances. Unable to differentiate the hypothesis, deciding on which test to use,one-tailed or two-tailed. **COMP**: They expect very often an exact and precise answer to the problems. **INT**: Unable to relate the results with the hypothesis being tested. Not sure of what conclusions they were coming to, whether to accept or reject.
18 years	L8	**PKNOW**: Students are so used to doing Maths and always thought of statistics as part of it. **ATT**: Some students just don't work hard enough. **TCF**: "The fact that I am a lousy teacher makes it more difficult for the students."	DK-1	**COMP**: Difficulties in handling technical questions without assistance from tutors or lecturers. **INT**: Unable to relate and interpret into some practical context.

DK-1: Question not asked of respondent
DK-2: Questions asked but not answered (strayed from question)

Table continue on next page.

15 years	L9	**PKNOW**: Most of them have not been exposed to statistics before. **ABLE:** Students are not mathematically inclined. Cannot relate statistics to their own subject area. **ATT:** They didn't like the subject because of the way it was taught. It was as they say, "...thrown at them-the mathematics, the formula"	**TF2:** "I think time has got a lot to do with it. Not enough time to make the teaching more memorable...less use of practicals and examples." **PRESS:** Have to complete the syllabus within the specified time. Tends to rush teaching towards the end.	**THEO:** How to measure variability? Cannot relate from a sample to a population. Cannot differentiate between test statistics value and p-value. The notion of test statistics and decision rule. INT: Cannot distinguish variations between two groups. Could not explain why one s.d. value is bigger or smaller than the other.
20 years	L11	**PKNOW**: They meet very little statistics in other part of the course They find statistics vague and less clear cut than the main stream mathematics. **ATT**: There maybe an earlier resistance to the subject because its different from the sort of maths they have experienced before. They don't have the stamina or the inclination to tackle lots or problems Don't see the relevance of statistics in their own discipline.	**TF2:** Lecturer tend to get bogged down with the traditional method to introduce difficult concepts and never actually given the time constraints to do more work that are more relevant to students. **PRESS**: Having to show the relevance of statistics in their own discipline. Lecturer has to know more about students area of discipline before presenting any examples related to their area.	**THEO**: Difficulty in applying rules of probability to problems. Cannot distinguish between density function and probability function. **COMP**: Unable to understand formulas and translating it into appropriate arithmetical operations. Resistance to the idea of using computers to do simulations. Basic level of numeracy-no feel for magnitude of numbers, associating units with measurement skills, when is it appropriate to do operations on certain type of numbers. **INT**: Do not understand the practical value of the result at the end. Unable to interpret result in a practical way.
15 years	L12	**ABLE:** Mathematics involved is a little too complicated for some of them, unable to cope especially those who are mathematically less able.	**PRESS**: Teacher find probability concept the hardest to teach. She tends to shy away from that and do as little as she can.	**THEO**: Distinction between sample mean and population mean Significant Test-able to follow instruction but fail to understand the theory behind it. **COMP**: In some cases, they find the formula a little too abstract to understand and to apply to the problems.

Table continue on next page,

20 years	CL14	**TF1:** "We need the time for them to sink in...there's too much in the syllabus moving along."	**TF2:** There is not much leeway for experimenting and trying new ideas. **PRESS**: Time are enhanced with modules. They have the assessment made out and standard procedures to follow...tend to teach along these lines.	**THEO**: Difficulty in distinguishing between different types of prob. dist. and in which circumstances they should be using. Conditional probability-in which circumstances to use. **COMP**: Difficult to appreciate the concept of variation, regression & correlation. They find formulas rather daunting, they tend to make mistakes frequently trying to calculate these things.
21-30 years				
30 years	T13	**TCF**: Ideas not properly build up. Pupils not given the right explanation of what to measure. Terms not explained carefully. Concept teaching very disjointed. Pupils are mainly given recipe for doing something. **PRAC:** Pupils are not given the chance to explore their own mind and initiatives through practicals, experiment and investigative approach. **BKNOW**: They need much more background knowledge to introduce to what sort of things they are thinking about.	DK-1	**THEO**: Unable to decide which test to use, one-tailed or two-tailed and how to use them In the statistical tables-deciding on the significance level and what it means. They do not understand where it's coming from and whether they are testing properly. Pdf- no understanding of what pdf is- the relations between histograms and the relative frequencies.
30 years	L1	**PRAC:** They ought to be told about their professional job, about having to make a decision, that they ought to deal with real data by adopting statistical techniques. What's missing is students are seldom exposed to real data collection, something that could get them motivated.	DK-1	**COMP:** Students do things mechanically without understanding for example they use computer by following instructions with no basis of understanding of the motive of the exercise.

DK-1: Question not asked of respondent

Table continue on next page,

28 years	L3	**PRAC:** The whole statistical approach, the statistical way of thinking is a new and foreign idea and until they have experienced quite a lot of practical use of this way of thinking, it's not easy for them to get a firm grip of what is generally going on. **ABLE:** Some students have limited mathematical background and ability to handle figures and manipulation of numbers. They could not accept ideas that do not fall within the scope that they can handle. **CD:** "Some of these concepts are quite sophisticated and they are difficult for us all to keep up clearing our minds even the most experience don't, so it's not surprising students find it to be difficult as well."	DK-1	**THEO:** Ideas about probability theory-the concept of events, the intersection of events, conditional probability-in which circumstances to use it. Fail to fully grasp the ideas behind the concepts of statistical inference-Confidence limits, Sampling dist. of means, Hypotheis testing .
29 years	L4	**ABLE:** "All concepts are in the sense difficult for the students to get a tremendous grasp of and it just depends on the level on which you expect one to gain expertise." **TF1:** "You can't expect the students to have a proper understanding of the concepts straight away. It is something that they gradually gain." **CD:** Some concepts are primarily quite difficult for the students to understand.	DK-1	**THEO:** How to generate probability function. Probability theory-conditional probability, when to use it, in what type of situations. **INT:** Students can relate the concept to the example but what they have difficulty is to relate it to a different example.
28 years	L7	**PKNOW:** They are very lacking in statistical knowledge in primary and secondary years. **ABLE:** They find it difficult when comes to looking at sophisticated testing procedures. They were immediately put off by the mathematical notations.	**TF2:** Time factor is often put forward as a problem but this can be tackled from the beginning through proper time management and planning .	**THEO:** Significance testing procedure-the setting up of hypothesis. The notion of rejecting and accepting the null. Confidence limits - the idea of interval estimate, what does it measure.

DK-1 = Questions not asked of respondent

MATRIX DISPLAY 2 of Conceptual Framework 2

Level of Experience	Subject	Methods used to teach concepts	Purpose of using the Methods	Implicationns of the Methods on Pupils and Students
10 years or less				
3 years	T2	1. Worked Examples 2. Tutorial-problem sheets	1. N/A 2. To let pupils do questions using techniques taught by teachers.	1. N/A 2. To assist teacher in knowing the areas that pupils find difficult. With teachers' assistance, pupils began to understand better. Opportunity to discuss indivi-dually about their difficulties with teacher
8 years	T4	1. Tutorial-problem solving through small group discussion 2. Investigative Approach	1. Practice-get more experience of handling problems. 2. To relate concepts to what they are doing, say through projects and practicals. To improve reasoning qualities and reasoning powers.	1. They have a much better working knowledge of the topic, worked on it for several hours and get their ideas together 2. Reasoning qualities are likely to improve with more involvement in practical work.
7 years	T5	1. Through worked examples 2. Problem solving through small group discussion	1. To illustrate concepts in worked examples. 2. To encourage pupils to talk about their assignments and investigations.	1. Pupils can understand concepts better with good explanations and illustration of examples. 2. They find it helpful to talk about problem situation among themselves-promotes interaction.
2 years	T6	1. Through examples and picture illustration. 2. Problem solving in small group discussion. 3. Lecture cum printed notes with blanks and gaps 4. Investigative Approach	1. To introduce concept, to illustrate different ways to approach problems. 2. To encourage discussion-promote interaction between pupils and teacher. 3. Teacher go through printed notes where pupil will put in extra notes, formulas and examples. 4. To encourage and promote discussions-get them to think about the problem.	1. It takes quite a while for them to assimilate ideas but they eventually have some standing with them. 2. Pupils are quite receptive when ideas are thrown at them. They are willing to give their answers and discuss the situation. 3. Pupils become much more attentive and understand better. 4. Mixed reaction from pupils. With a little bit of guidance, they get to think and present their ideas on the problem-it does help their understanding.
1 year	T8	Basically chalk and talk	N/A	N/A
1 year	T10	1. Lecture-through example 2. Printed notes 3. Problem sheets	1. To introduce and emphasise concepts. 2. To provide detailed infor-mation about the topic. 3. N/A	1. Pupils prefer working through examples step by step. 2. They don't like it very well-tend not to be attentive. 3. Prefer the exercise to help them to understand.

2 years	T11	1. Traditional teaching-chalk and talk 2. Examples on board. 3. Do questions out of textbooks. 4. Tutorials-problem exercise.	1. N/A 2. N/A 3. To let pupils familiarise with questions, to see how questions are phrased and understand from there. 4. To help them with the exercise. To find out if they have any difficulties in under-standing certain concepts.	N/A to all
10 years	CL5	1. Basically teaching, not lecturing 2. Problem sheets	1. N/A 2. To let them work through the problem sheet-to find out if they have difficulties in understanding.	N/A to all
8 years	L10	1. Simulation-repeated sampling 2. Mainly lecture in a traditional way 3. Slowly changing to more practical work and compu-ter exercises.	1. Looking at the effect of the spread of different sample sizes leading to discussion of variation and std. deviation 2. N/A 3. To impart skills in handling real data and practical situation.	1. Students respond better-step by step through the whole questions. 2. N/A 3. Their learning approach changes. It is worthwhile at the end because students really appreciate doing something, seeing how it works out in new context rather than passively receiving information in the lecture.
3 years	CL13	Basically teaching-use of OHP and board.	To explain teaching notes and illustrate examples.	N/A
11-20 years				
15 years	L2	Start with a good example-propose a problem to the students.	N/A	Better understanding of the problem when more feedback received from them.
20 years	CL6	1. Talking through the concepts-give examples of situation on board. 2. Question and Answering situation 3. Investigative approach	1. "To present the concepts and ideas, putting it on the board as I go along." 2. To obtain feedback, to help teacher to find out which area they have difficulty with. 3. Students gather information and present it in the form of analysis and report.	1. "They can picture better what you are doing. It gives them a chance to sink in as I am writing." 2. It helps students to identify their own problems. Knowing which area they should put in more effort. 3. N/A

18 years	L8	1. Chalk and talk 2. Start with an example-Biology, maths area 3. Lecture notes-prepare in advance 4. Tutorial- small group discussion 5. Lab-practical work 6. Cooperative learning 7. CBL	2. Translate the example into statistics-introduce notations, perform test and interpret results. Less mathematical terms and definition. For mathematicians-same approach but more formal definitions and notations. 3. To provide more information about the topic. 4. Discuss questions which they have done in advance-let them discuss among themselves 6. Appropriate in project work 7. Used as revision purpose, not used during the lecture.	1. It depends very much on the audience. A lot of students actually enjoy the blackboard. Students can keep up as the material is absorbed into their mind. N/A for (2) - (5) 6. It helps in the learning of others. Weaker students may benefit from the strong ones. A lot of collaboration involved.
15 years	L9	1. Lecture 2. Tutorial 3. Use incomplete work-books 4. Labwork, Lab workbook	1. N/A 2. N/A 3. Students fill in gaps in workbook such as formula, worked examples, proofs 4. To reinforce some of the lecture materials using Minitab To get them do things in the lab that might help them understand what was done in the lectures.	1. N/A 2. It slows down teaching, however, it allows students to catch up with learning. 3. N/A 4. They like the hands on exercise, allow them to exchange ideas and discuss work among themselves
20 years	L11	1. Traditional lecture-chalk and talk 2. Printed notes and OHP materials 3. Use of examples 4. Introduce a little humour -cartoons, anecdotes 5. Tutorial bulit into lecture 6. Computer labs 7. Small group discussion	1. Deliver lecture materials. 2. N/A 3. To provide theories of application for statistics in their own discipline. 4. To lighten the atmosphere. To raise a general awareness about the misuse of statistics 5. To encourage student to talk about their difficulties in understanding concepts, to use staff as support activity rather than rely on them to provide the solution. 6. Work on practical example using Minitab. 7. Students to work on problems and collaborate with one another.	1. N/A for (1), (3), (4) & (6) 5. "Not many will come along in the way that you hope. They are usually gone before they met anything. Some resent a bit strong and not looking for you to interlude." 7. More successful as they get to do more on their own rather than listen.
20 years	L12	1. Traditional lecture-chalk and talk. 2. Formula and examples	1. N/A 2. To explain concept and what formula means relative to the concept, how calculation works using formula.	1. Students take a long time to write and not use abbreviations. Tendency to write everything. 2. N/A

20 years	CL14	Traditional lecture-chalk and talk.	Introduce the theory and then try to get appropriate examples from their field of interest.	N/A
21-30 years				
30 years	T13	1. Simulation using graphic calculator. 2.Theoretical explanation and problem solving - a combination of both. 3. Tutorial approach	1. To build up the theory behind it, to get them to see the ideas clearly from the very basic. To introduce things in investigative ways. 2. To provide understanding of the background. 3. Introduce the subject by putting examples on the board and getting them to fill in the next step.	N/A for (1) & (2) 3. A lot of interaction going on, obtained quite a lot of feedback from them. They have tried to grasp what is happening.
30 years	L1	1. Traditional lecture-chalk and talk. 2. OHP materials 3. Small group tutorials 4. Anecdotes	1. To deliver lecture materials. 2. Use OHP to explain examples. 3. N/A 4. To give examples of a practical use of statistics.	1. Students are comfortable with it-they have time to copy while listening to his lecture. When asked, they prefered it as well. 2. It didn't work-tutors are not trained to teach Pharmacy students. They tend to neglect what students need to know. Process is slow, takes a longer time for tutor to explain certain concepts and students take longer time to complete certain tasks.
29 years	L4	1. Lecture-for e.g. Probability concept, introduce it purely from the frequency approach, excessive ideas of sample space, lots of definitons. 2. Examples for every new concepts.	1. N/A 2. Every example has two parts. One is translating a story, a proper well-defined statistical questions. Two, being able to solve the question.	1. The very bright student paradoxically are far more motivated than the moderate students. Thus, the brighter ones should be able to cope. 2. The problem is a lot of students regard the story as being a different part of the series. So it's the english to the statistical questions. They find it difficult to disentangle the question story.
28 years	L7	1. Traditional lecture-chalk and talk. 2. Present ideas in a nice, simple way through examples.	1. To deliver lecture materials. 2. Make sure not to go too basic, avoiding too technical words. Examples used are link to ordinary situation, real data.	1. N/A 2. They understand it quite well

N/A: Information Not Available

MATRIX DISPLAY 3 of Conceptual Framework 3

Level of Experience	Subject	Signpost for students' misunderstanding of concepts	Signpost for students' understanding of concepts	Methods used to evaluste understanding/misunderstanding
10 years or less				
3 years	T2	• From their work-they will admit to not being able to do the questions, not understanding the examples. • Not following the procedure when comes to testing the hypothesis. • Unable to interpret ideas and results from the test. • Unable to interpret diagrams.	N/A	•Informally, from the work that they produced. • From the way they asked questions.
8 years	T4	• From the practicals-could not interpret results properly.	• They do better in questions on certain topics. • Higher marks achieved from test and projects.	• Informally, rely on them to tell of their problems in understanding certain concepts.
7 years	T5	•They will admit to not being able to do the problems.	N/A	• Class is small-usually rely on them to tell me. • Do not do much formal testing.
2 years	T6	• Poor preliminary results.	• Good preliminary results.	• From preliminary exam. • Do not do much formal testing due to time constraint.
1 year	T8	• "If they failed means they don't do well and that they don't under-stand the question."	•"If they passed means they do well and that they do understand the questions." • Correct solutions are performed based on their understanding of the questions. They have to understand before they can actually do it	• Basically from preliminary examination.
1 year	T10	• Not anwering the questions that was posed to them. •They will tell which part they are having difficulties in understanding.	• By producing good test results. •They used the correct method to solve the problem.	• Homework exercises. • Class test and prelim. exam. •Talking to them in groups and individually.
2 years	T11	• Not following the right procedure in testing hypothesis. •Wrong formula use in the calculations. • Fail to interpret results properly.	N/A	• Homework exercises-collected and marked

Table continue on next page.

10 years	CL5	• The feedback with the type of problems they have has determined how much they have taken in.	• Pass each performance criteria as a result of some basic understanding of the concept.	• Tutorial worksheets. • Assessment from each outcome and performance criteria (HND, HNC)
8 years	L10	• From group and individual discussion regarding notes, lectures and problem exercises. "I had a pretty good impression that they really don't understand what was going on."	• From the project report-they understood some fairly good concepts. • Practical session-a vast majority of the students seem to have grasped the idea of what they are suppose to do each time.	• Through instinct of looking around and of hearing what students had to say. • Projects and practicals.
3 years	CL13	• Looking at solutions produced for questions-unable to perform calculation correctly, methods not used appropriately.	• Pass each performance criteria due to some basic understanding of concepts.	• Tutorial worksheets. • Assessment from each outcome and performance criteria (HND, HNC)
11 - 20 years				
15 years	L2	• Students tried to avoid questions on theoretical proofs and those that require explanation and interpretation.	N/A	• Exam results. • Rely on them to tell.
20 years	CL6	N/A	N/A	• "Tutorial exercise-give examples, asked them to follow through several steps. Then I would give some questions and see how they do it."
18 years	L8	• "From the way they respond to the question, I can get the understanding of who knows what's going on and who doesn't." • From the way they ask certain questions indicates that they are not understanding.	N/A	• Through written exams, report on project and practical work. • Tutorial worksheets-To see how well the students do by the quality of the materials they submit. • Informal conversation with students.
15 years	L9	N/A	N/A	• Exam results. • Students' feedback-from the questions they asked and the work they do.
20 years	L11	N/A	• Can see some sparks in the students attitude towards the subject, for e.g. nodding their heads- a sign of following and understanding what was taught.	• Tutorial work-exercise.

Table continue on next page.

18 years	L12	N/A	N/A	• Project report •Informal conversation with students.
20 years	CL14	•From the solutions of exercise that they do, estimate from there that they find it difficult to grasp certain ideas.	N/A	• Tutorial work
21 - 30 years				
30 years	T13	N/A	•"Once they are able to do the work with a variety of problems, you know somehow they have understood it." • "If they can use it in a practical situation, then they must have understood it."	• Exercise worksheets • Practical assessment
30 years	L1	•Fail to use appropriate statistical techniques. • Produced wrong results in certain problem exercise.	• Use of correct statistical techniques. • Able to take problem context and translating it into the statistical problems, going through the technical exercises. • Able to produce correct results and make the right conclusions. • Able to make inferences, a decision in their professional life. • Able to use results in decision making.	• Computer exercise. • Theoretical standard examination questions. • Projects on different themes.
29 years	L4	N/A	N/A	• Records of marks in the tutorial-give an idea how they are performing, which area they find difficult. •Lecturer marked all papers-get a very good idea from there.

N/A: Information Not Available

Table 8.5: Content analytic summary on the implications of methods used on pupils and students

Methods	Implication Items	Teachers' perception	Lecturers' perception
Worked Examples	Provide Assistance	T2: To assist teacher in knowing the areas that pupils find difficult. With teachers' assistance, pupils began to understand better.	
	Promote Discussion and Interaction	T2: Opportunity to discuss individually about their difficulties with teacher.	
	Enthusiasm	T5: They like doing the exercise but tend to rush through the calculation without needing to understand.	
	Promote Understanding	T6: It takes quite a while for them to assimilate the ideas but they eventually have some standing with them.	L2: Better understanding of the problem. L4: The problem is a lot of students regard the story as being a different part of the series. They find it difficult to disentangle the question story.
Problem Solving through small group discussion	Promote Understanding	T4: They have a much better working knowledge of the topic- worked on it for several hours get their ideas together.	L1: It didn't work-tutors are not trained to teach Pharmacy students. They tend to neglect what students need to know. Process is slow, takes a longer time for tutor to explain to students certain concepts and students take longer time to complete certain task. L11: More successful as they get to do more on their own rather than listening passively.
	Promote Discussion and Interaction	T5: They find it helpful to talk about problem situation among themselves-more interaction.	
	Enthusiasm	T6: People are quite receptive when ideas are thrown at them. They are willing to give their answers and discuss the situation.	

Table continue on next page.

Investigative Approach	Reasoning Quality Promote Understanding	T4: To improve their reasoning qualities, to develop their reasoning power. T6: Mixed reaction from pupils. With a little guidance, they get to think and present their ideas on the problem-it does help their understanding.	N/A
Lecturing/Teaching the traditional way	Adapting	N/A	CL6: They can picture better what you are doing. It gives them a chance to sink in as you are writing. L8: It depends very much on the audience. A lot of them actually enjoy the blackboard. Students can keep up as the material is absorbed into their minds. L12: Students take a long time to write and not use abbreviations They tend to write everything. L1: Students are comfortable with it-they have time to copy while listening to his lecture. They prefer it as well.
Printed Notes and Workbooks	Provide Assistance Promote Understanding	T6: Pupils become much more attentive. T10: They don't like it very well-tend not to be attentive. T6: Better understanding of the concepts	L11: Could not put in more examples and illustrations due to lack of preparation time.
Practicals and Projects	Promote Learning Promote Discussion and Interaction	N/A	L10: It does make a difference to how they are learning. It's worth at the end because students really appreciate doing something, seeing how it works out in new context rather than passively receiving information in the lecture. L9: They like the hands on exercise, allowing them to discuss among themselves.
Simulations - Computers, Graphic Calculators	Promote Learning	N/A	L10: Students respond better-step by step through the whole question.

Table continue from previous page,

Others-CBL,Q/A, Coopearative Learning, Anecdotes	Promote Discussion and Interaction	N/A	L8: Cooperative Learning-it helps in the learning of others. Weaker students may benefit from the strong ones.- a lot of collaboration CL6: Q/A-They learn better, a lot of feedback and interaction between students and lecturers. L1: Anecdotes-Getting anecdotal during the teaching at some parti-cular point to include some examples from the actual research. to get students to understand the concept relative to the actual thing.

N/A Information Not Available

Table 8.6: Content analytic summary of methods used to evaluate pupils' and students' understanding and misunderstanding of the concepts

* ITEMS	Teachers' View	Lecturers' View
Informal conversation, rely on students to tell (INFORM)	T2: Informally, from the work that they produced, they will tell which part they are not understanding. T4: Rely on them to tell T10: Talking to them in groups and individually.	L2: Rely on students to tell of their difficulties in understanding. L8: Informal conversations with students. L11: Talking to students in small groups.
From the nature of their question (ASK)	T2: From the way they ask questions, teacher had the impression of whether they are understanding or not.	L9: From the questions they asked and the work they do gives the lecturer some idea of how much the students have understood the concepts. L12: From the questions they asked.
Test and Examination Results (EXAMS)	T6: From preliminary exam T8: Basically, from preliminary exam T10: Class test and preliminary exam	CL5: Assessment from each outcome and performance criteria (HND, HNC) CL13: Assessment from each outcome and performance criteria (HND, HNC) L2: End of term examination L8: Through written exam L9: End of term exam L11: Course assessment - tests, exams L1:Theoretical standard exam questions L4:Record of marks in the exams
Homework Exercises (HOMEWK)	T10: From homework exercises-collected and marked T11: Homework exercises- collected and marked	L9: From the homework that they do T13: From exercise worksheets given at the end of the lecture-homework exercise

Tutorial Worksheet Exercises (TUTOR)	N/A	CL5: From tutorial worksheet-discuss during the tutorial CL13: Tutorial worksheets CL6: Tutorial exercise-give examples and asked them to follow through several steps. Then give them some questions and tried to get them to do the same way. L8: Tutorial work-could see how well students do by the quality of the materials they submit. CL14: Discussion of homework exercise during tutorial-get them to present their answers. L4: Records of marks during the tutorial-gives an idea how they are performing and which area they find difficult.
Students' Reaction (REAC)	N/A	L10: Through instinct of looking and hearing around what students had to say.
Projects, Practicals and Computer Exercises (PRACS)	T13: Practial assessment	L10: Assessment of projects and practicals L12: From project reports L1: Computer exercises and assessment of projects on different themes

* Methods used to evaluate pupils' and students' understanding/misunderstanding of the concepts.
N/A Information Not Available

Table 8.7 : Content analytic summary of teachers' and lecturers' views on the signpost for pupils' and students' misunderstanding of the concepts

Signpost Items	Teachers' View	Lecturers' View
Computational Procedures (COMPROC)	T2: Not following the procedures when comes to testing the hypothesis. T11: Not following the right procedure in significance testing. Wrong formula used for calculations.	CL13: Unable to perform calculations correctly. methods not used appropriately. L10: From their solutions to exercises- calculations not performed according to instructions. L2: Students avoid questions on proofs and meticulous calculations. CL14: From the working in the exercise, the lecturer could estimate how difficult they are finding it, whether or not they are grasping the ideas. L1: Do not use correct statistical techniques. Unable to take problem context and translating it into the statistical problems: failure to go through the technical exercises correctly.

Feedback on Questions Asked (FEED)	N/A	L10: From their solutions to exercises, the lecturer had a good impression that they don't understand what was going on. CL5: The feedback with the type of problems they have determine how much they have taken in. CL13: Looking at the solutions they produced for the questions-lecturer had a good idea of which area they are not understanding. L8: From the way they responded to certain questions-not satisfying. CL14: Not responding correctly to questions asked.
Interpretation of Results (INTRES)	T2: Unable to interpret ideas and results from the test. Unable to interpret graphs and diagrams correctly. T4: From the practicals-could not interpret results properly. T11: Unable to interpret results properly.	L1: Could not produce results, interpretation and conclusions as requested.
Test and Examination Results (EXAMR)	T6: Poor preliminary results. T8: Failing means they don't do well, that they don't understand what the question was asking.	N/A
Admitting to Problems (ADMIT)	T2: They will admit to not being able to do questions and not understanding what was taught. T5: They will admit to not being able to do the problems during the tutorial. T10: They will tell which part they are having difficulties in understanding.	N/A
Avoiding Questions (AVOID)	T10: Avoiding questions that they feel they could not handle-no effort in trying.	L2: They admitted to avoiding questions that require extensive calculations, explanation and interpretation.
Relate to Experimental and Practical situations (PRAC)	N/A	L1: Unable to make inferences, a decision in their professional life. Unable to use results in decision making.

N/A Information Not Available.

Table 8.8: Content analytic summary of teachers' and lecturers' views on the signpost for pupils' and students' understanding of the concepts

Signpost Items	Teachers' View	Lecturers' View
Computational Procedures (COMPROC)	T10: Correct methods used to solve problems T13: "Once they are able to do the work with a variety of problem exercises, you know somehow they have understood it."	CL14: From the work that they do in the exercise-correct use of calculations and manipulation of numbers. L1: Correct use of statistical techniques. Able to take problem context and translating it into the statistical problems, going through the technical exercises.
Feedback on Questions Asked (FEED)	T4: They do better in questions on certain topics. T8: Correct solutions are performed based on their understanding of the questions.	L8: "From the way they responded to the questions correctly, I can get the understanding that they are able to grasp the idea."
Interpretation of Results (INTRES)	N/A	L1: Able to produce results and interpreting it correctly.
Test and Examination Results (EXAMRES)	T4: Higher marks achieved from tests and projects. T6: Good preliminary results. T8: If they pass means they do well and that they do understand some of the questions. They have to understand before they can produced the correct solutions. T10: Produced good test results.	CL5: Pass each performance criteria as a result of some basic understanding of the concepts.
Better Response (BETRES)	N/A	L11: Can see some sparks in the students' reaction during the teaching.
Relate to Experimental and Practical situations (PRAC)	T13: If they can use it in the practical situation, then they must have understood it.	L10: Practical session-a vast majority of the students seem to have a pretty good idea of what they are trying to do each time. L1: Able to make inferences, a decision in their professional life.

N/A Information Not Available

Measure of distances between codes in Matrix Display 1, graphical displays of MDS and cluster analysis using Splus

---

```
> x_matrix (scan ("matdat1"), byrow=T, ncol=17)
> dimnames (x) [[2]]<-list(c ("PKNOW", "ABLE", "ATT", "TF1", "CD", "TCF","LCON1", PRAC", "BKNOW", "LEX", "LCON2", "PRESS", "TF2", "RES","THEO", "COMP", "INT"
> openlook ( )
> par (mfrow=c(2,1))
```
note: calculation of distances between codes
```
> z < - dist (t (x), metric="binary")
```
note: to produce complete linkage dendrogram
```
> plclust (hclust (z, method="com"), labels=dimnames (x) [[2]] )
```
note: to produce 2-dimensional solution of MDS
```
> y_cmdscale (z)
> plot (y, type= 'n')
> text (y, dimnames (x) [[2]] )
```