CHAPTER 5

OVERVIEW - STUDY 1

In the previous chapter the results were considered in detail for each of the four focus areas. These results are now combined to present an analysis of the overall level of understanding exhibited by the students' responses, and a comparison of the difficulty of the questions. This is done in an attempt to provide a general description of the level of students' understanding of statistics and to reinforce the levels suggested in the hierarchy in Chapter 4.

There are four main sections in this chapter. First, in the **Overall Student Performance** section, the level of student understanding in statistics, in general, is considered rather than in specific focus areas. The factors of academic year, mathematical ability and gender were found to have various effects on understanding over the four focus areas and the results are now analyzed to determine whether there is an influence on overall understanding. Next, **Question Difficulty**, compares the statistical understanding required to progress through the various levels of the four questions. This is done to determine which, if any, of the questions were more difficult. **Level Thresholds**, the third section, compares the amount of statistical understanding required to produce a response at a certain level. This is necessary to help justify the cycles of levels that were suggested in Chapter 4, and also to present a profile of the 'average' student. The final section, **Conclusion**, summarizes the findings of Study 1.

Overall Student Performance

To investigate overall student performance, the levels for each question were processed using QUEST. Due to the low number of students in Levels 0 and 8 being unacceptable to the analysis package, all Level 0 responses were recoded as Level 1 and all Level 8 responses were recoded as Level 7. Effectively, this gave the range in Levels from 1 to 7 but this was considered not to detract from the overall results. The item consistency for the test analysis was 0.7 which is considered sufficiently high (Vine, 1994) to assume that the items were all consistent in testing the same underlying construct, in this case, statistical understanding. The fit statistics, for the question estimates are presented in Figure J.1 and for the student estimates in Figure J.2 (Appendix J). Both infit mean squares are not significantly different from 1, which means that the data fitted the Rasch model well. The infit mean square map for the four questions is shown in Figure J.3. All questions showed values within the acceptable limits indicated by the dotted lines produced by the QUEST package.

This implies that the parameters for each question fitted sufficiently well into the model to contribute to the estimate of understanding.

As a result of the QUEST analysis, each student was given a score which was an indication of overall performance over the four focus areas. This estimate was taken as a measure of a student's understanding of statistics, the higher the estimate, the greater the understanding. These estimates are now used to determine whether the factors academic year, mathematical ability and gender influence the level of a student's understanding of statistics.

The spread of estimates is shown in Figure 5.1. Each X represents the estimate for one student and the - - - - - line represents the mean estimate for all students. The lowest estimate was -0.99 and was recorded by some students in Years 7 and 8. The highest estimate was 2.16 for a Year 12 student. This was exceptionally high compared to the next highest score which was 1.83. Only 12 of the 180 students attained scores of 1 or greater. The overall mean estimate was 0.2 and the spread of the estimates is reasonably well balanced, except for the few high flyers mentioned earlier.

2.2	X
2.1	
2.0	
1.9	
1.8	XX
1.7	
1.6	XXX
1.5	
1.4	X
1.3	
1.2	X
1.1	
1.0	XXXX
0.9	XXXXXXX
0.8	XXXXXXXX
0.7	
0.6	XXXXXXXXXXXXXXXXX
0.5	XXXXX
0.4	XXXXXXXXX
0.3	XXXXXXXXXXXXXXXXXX
0.2	XXXXXXXXXXXX -
0.1	XXXXXXXXXXXXXXX
0.0	XXXXXXXXXXXXXXXXXXXXXXXXXX
-0.1	XXXXXXX
-0.2	XXXXXXXXXXXXXX
-0.3	XXXXXXXXXXXXXX
-0.4	XXXXX
-0.5	XXXX
-0.6	XXX
-0.7	
-0.8	X
-0.9	
-1.0	XXX

Figure 5.1 - Estimates of Understanding

An analysis of variance was performed on these estimates to test the overall influence of the three factors, academic year, mathematical ability and gender on statistical understanding. The results of the analysis are shown in Table 5.1. Very significant results for all three main effects indicate that the level of understanding exhibited by a student depends on his or her academic year, mathematical ability and gender. All interactions are not significant which suggests that, whatever the variation within the main effects, it is the same at all levels of the other factors. For example, whatever the difference between the level of understanding for males and females it is the same irrespective of their academic year or mathematical ability. These main effect differences are investigated further below.

Table 5.1 - Estimates of Understanding - Analysis Of Variance

Source	S.S.	D.F.	M.S.	F-Ratio	Prob.
Year(Y)	11.222	5	2.244	12.289	0.000
M.Ability(M)	4.095	2	2.048	11.211	0.000
Gender(G)	2.328	1	2.328	12.746	0.000
Y x M	1.165	10	0.116	0.638	0.780
Y x G	0.051	5	0.010	0.056	0.998
M x G	0.079	2	0.039	0.215	0.807
Y x M x G	2.078	10	0.208	1.138	0.338
Error	26.299	144	0.183		

The estimate means for each level of the main effects are shown in Table 5.2. The differences between pairs of these means, and the comparison probabilities for the Scheffe Test of these differences, are presented in Appendix I. Tables I.1 and I.2 show the mean differences and comparison probabilities for the statistical comparison of the level of understanding for the various academic years. Similarly, Tables I.3 and I.4 show information for the comparison of the mean understanding for each level of mathematical ability, and Tables I.5 and I.6 allow for gender mean comparisons.

From Table 5.2, it is apparent that the estimate means increase with increasing academic year, as would be expected. A graphical representation of these results, in Figure J.4 in Appendix J, helps to demonstrate the mean and spread of the estimates for each of the academic years. Tables I.1 and I.2 show that the academic year main effects become more apparent with increasing academic year. The means for the estimates of understanding for the first four academic Years, 7, 8, 9 and 10 are very similar and none of the pairwise comparisons were significant. The Year 11 estimate mean is significantly different to that of

Year 7, and the Year 12 estimate mean is significantly different to that of all other years. Hence, although there is a significant difference between the academic years in terms of understanding, the difference between successive years is mostly not significant. Most years of secondary schooling elapse for a student before there is a significant change in statistical understanding. The most significant change between successive academic years takes place between Years 11 and 12.

Table 5.2 - Estimates of Understanding - Main Effect Means

Year 7	-0.13
Year 8	0.05
Year 9	0.10
Year 10	0.17
Year 11	0.37
Year 12	0.65
Low Ability	-0.00
Middle Ability	0.26
High Ability	0.35
Female	0.31
Male	0.09

Referring again to Table 5.2, as mathematical ability increases the estimate means also increase, as expected. This information is represented graphically in Figure J.5 which demonstrates the greatest spread of understanding estimates for the students with low mathematical ability. Tables I.3 and I.4 show that the change in mathematical ability becomes less apparent with increasing mathematical ability. The mean for the low mathematical ability students was significantly different to that for the middle and high mathematical ability students. However, there was not a significant difference between the means for the middle and high mathematical ability students. This indicates that students with low mathematical ability show much less statistical understanding, on average, than those students with middle or high mathematical ability.

The average estimate for females, in Table 5.2, is considerably higher than for males. The comparison of the estimate means for gender, in Tables I.5 and I.6, shows that the difference is highly significant. This information is represented graphically in Figure J.6, which also shows a greater spread of estimates for the females than the males. Hence, females show both a greater range and a higher average statistical understanding than males. This may be partially accounted for by the form of written test, as female students may be more inclined to put more detail into written answers than male students.

All the factors, academic year, mathematical ability and gender influence the level of a student's statistical understanding. This was evident to varying degrees in Chapter 4 for the four focus areas and has now been demonstrated for overall statistical understanding. Although there is a slight increase in understanding with academic year it requires most of the secondary school experience before this becomes apparent. The lack of understanding of lower mathematical ability students suggests that care needs to be taken as to what is expected of these students when they are being required to process statistically. The poorer understanding exhibited by males may be due to the written form of the test and warrants further investigation, perhaps by interview, to determine whether males are able to improve on their performance when responding verbally. The levels of difficulty of the questions are now compared, as a means of reinforcing the steps between the various stages in the modes and cycles suggested in the hierarchies in Chapter 4.

Question Difficulty

The overall statistical understanding estimates calculated for each student made it possible to compare the questions, to determine whether students found some more difficult than others. In this section, the overall difficulty of the questions are compared first. Then, the difficulty is studied in more detail, by considering the ease with which students are able to move from one level to the next within a question.

The difficulty factor for each question is shown in Table 5.3, the higher the value the more difficult the question. Question 1 Part I was found to be the most difficult for the students. This perhaps reflects the lack of experience with sampling for secondary students. Question 1 Part II, as expected, was found to be the easiest as this is the question where some prompting was given. In the other three questions Part I, where the data were presented in raw form, was easier for the students to complete than Part II. However, in Question 4 the level of difficulty is almost the same for both parts. These results suggest that students may find statistical processing more difficult when the data are presented as a graph.

To analyze this difficulty in more detail it is possible to consider the degree of difficulty, or tau value, in moving from one level to the next in each question. The tau values are presented numerically in Table I.7 (Appendix I). Figures 5.2 to 5.5 present these values graphically for each of the four questions. Of particular interest, for each question, are the difficulties in moving into Level 3 and Level 6, as these are the stages where there are changes of mode, or cycle within mode, as described in Chapter 4.

Table 5.3 - Question Difficulty

Question 1	**Part I**	0.41
	Part II	-0.43
Question 2	**Part I**	-0.03
	Part II	0.24
Question 3	**Part I**	-0.21
	Part II	-0.04
Question 4	**Part I**	-0.01
	Part II	0.05

For Question 1, there is a similar pattern of tau values for Part I and Part II, except for the step from Level 2 to Level 3. The graph in Figure 5.2 shows that overall students found the step from one level to the next easier in Part II when prompt information had been given. However, it was far more difficult for a student to step from Level 2 to Level 3 on Part II than on Part I. This shows that, even with prompted information, students found it very difficult to begin considering the details of data collection. However, once begun, it was easier to go on to consider the data collection process in more detail. The difficulty of the step into Level 6 was reasonably high and this was the same for both parts of the question. This indicates that once data collection details have been considered, there was no advantage or disadvantage gained by being prompted with some earlier details of the sampling method.

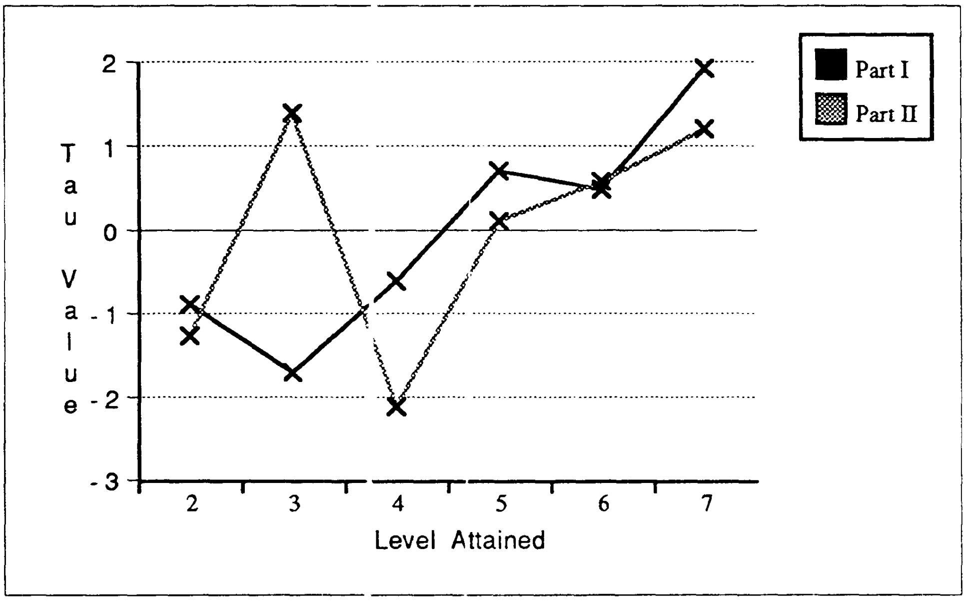

Figure 5.2 - Difficulty Moving Between Levels - Question 1

The graph in Figure 5.3 shows very similar difficulty for the steps between levels for Part I and Part II of Question 2. At all except one step, the tau value for Part II is slightly higher suggesting that when the data are presented as a graph it is slightly harder for a student to produce a response at the next level. The steps into Level 3 and Level 6 are amongst the highest values suggesting that it is more difficult for students to enter either of the two cycles in the concrete-symbolic mode than to move within the cycle. The only higher value is from Level 3 to Level 4 in Part II of the question. The easiest steps appear to be into Level 2 and Level 5. The low difficulty of attaining an entry to Level 5 suggests, that once students have a basic understanding of what a table or graph is presenting, they are able to combine all factors involved with relative ease.

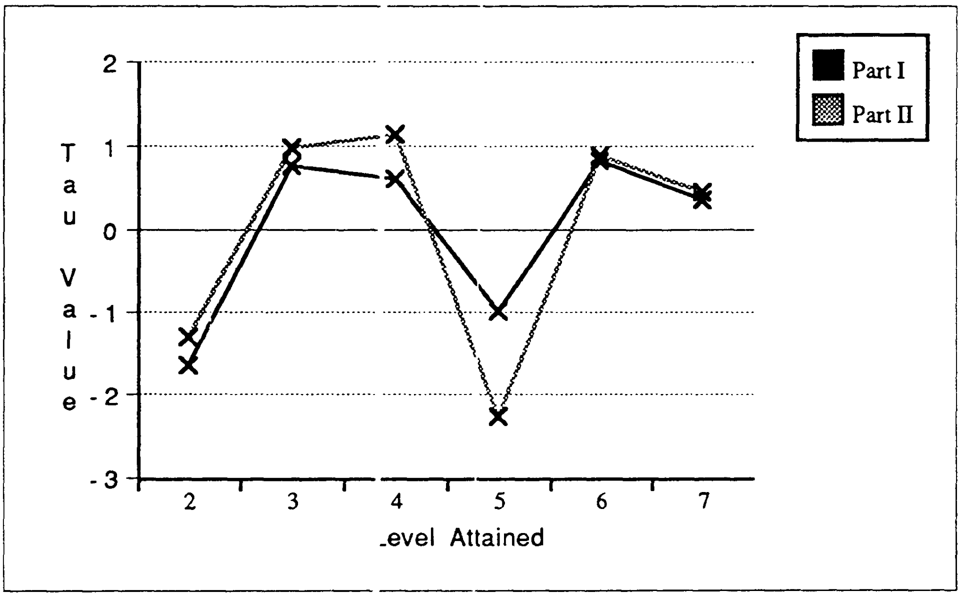

Figure 5.3 - Difficulty Moving Between Levels - Question 2

It appears, from Figure 5.4, that for Question 3, a step to the next level is slightly easier at the lower levels when the data are presented as a graph. However, there is a big difference at the point of entry into Level 6, where it appears to be far more difficult for a student if the information is presented as a graph. The easiest step in both parts is from Level 3 to Level 4, which is within the first cycle of the concrete-symbolic mode. The step into the beginning of this cycle, Level 3, is quite difficult. The only steps more difficult, being into Level 6 and 7 for Part II and Level 7 for Part I. In fact, it appears to be very difficult for students to make that step into the second cycle of the concrete-symbolic mode, using recognizable statistics to summarize data when the information is presented as a graph. They find it much easier to make that step when the information is presented in raw form which is probably a reflection of the type of experiences which are presented most commonly to them in the learning environment at school

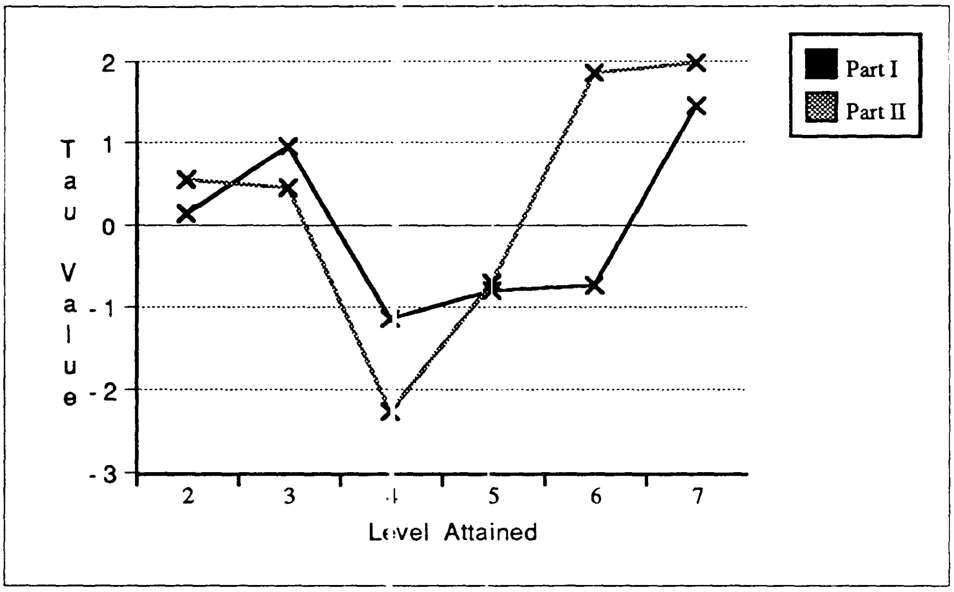

Figure 5.4 - Difficulty Moving Between Levels - Question 3

The tau values in Figure 5.5 suggest that there is very little difference between the difficulty of attaining certain levels for Part I and Part II of Question 4. The easiest steps appear to be moving into the latter part of the first cycle in the concrete-symbolic mode. Once students have been able to recognize a simple pattern in the data, they find it relatively easy to describe that pattern or make predictions in simple terms.

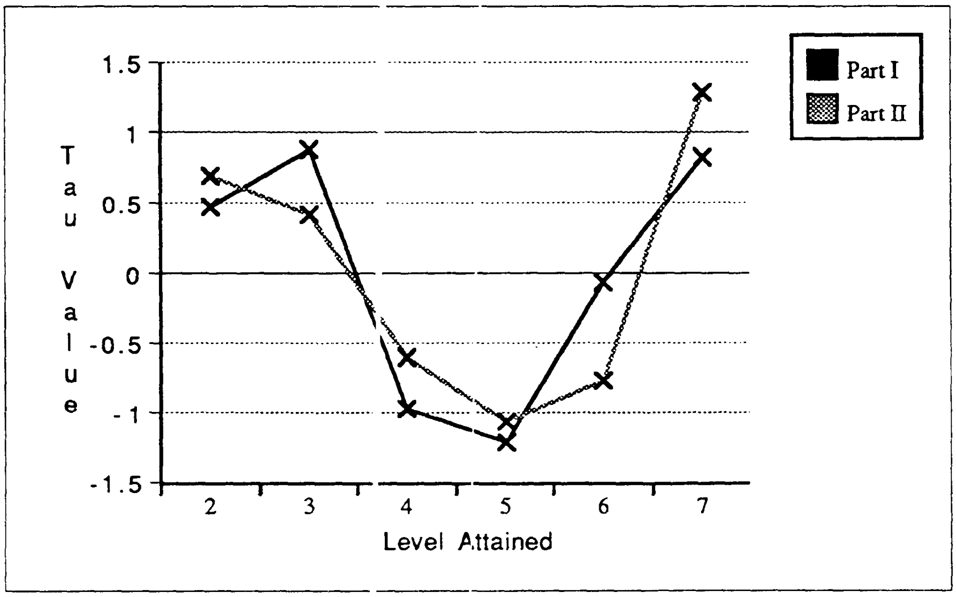

Figure 5.5 - Difficulty Moving Between Levels - Question 4

However, the high tau values for the steps into Level 2 and Level 3 show that students found it very difficult to get started in the process of interpretation and inference. The step into the second cycle of the concrete-symbolic mode, Level 6, was not as difficult as getting started but it was somewhat more difficult than staying in the first cycle. Again, students found it very difficult to attain the detail that was necessary to produce a response at Level 7.

The investigation of question difficulty has shown that similar patterns of tau values were exhibited for Parts I and II within each of the questions but there was some variation between questions. Every question, with the exception of the prompted part of Question 1, had a high difficulty for the step from Level 2 to Level 3 which is the stage at which students are moving from the ikonic to the concrete-symbolic modes. The step into Level 6, the second cycle of the concrete-symbolic stage was reasonably difficult for most of the questions, although it was surprisingly easy for Question 4. This possibly suggests that although the fourth question was found to be difficult, once students progressed into the concrete-symbolic mode, it was not such a major step to progress to the stage where more sophisticated statistical descriptions are being used for interpretation and inference.

As well as the demarkation of the various modes and cycles, the estimates of understanding can be used to consolidate the structure of the levels within the cycles and modes. This is considered in the next section.

Level Thresholds

Threshold values, produced by the QUEST analysis for the various levels in each question, are used to justify the levels within the cycles described in Chapter 4 and to create a profile for an 'average' student. These values are estimates of the score a student would need to attain a 50% chance of having his or her response coded at that level. For example, the threshold value of 0.15, for Level 5 of Question 2 Part I, means that a response from a student with an estimate of 0.15 has a 50% chance of being coded at Level 5 for that particular question. Table I.8 contains the threshold estimates for the four questions. These results are presented graphically in Figures 5.6 to 5.9. No threshold value is shown for Level 1 because there is no information about performance at a level below this to allow estimation of the better understanding needed to be able to be coded as Level 1 (see Chapter 3).

Level Justification

Three levels were defined within each of the two cycles of the concrete-symbolic mode in Chapter 4. These corresponded to Levels 3, 4 and 5 within the first cycle, and Levels 6, 7 and 8 within the second. Threshold values are now used to justify these levels being arranged into cycles. It is anticipated that the threshold value for the levels within a cycle should be similar, indicating that once a student is able to function within a cycle there is not as great a change in the level of understanding to be able to progress from one level to the next within the cycle. Then, the threshold should increase to move into a new cycle.

Figure 5.6 shows that, for Question 1, there is a steady increase in the threshold value required to attain each of the levels. The threshold values for Part II, where some prompting was given, are consistently lower for Part I as may be expected. These levels are not necessarily supportive of the cycles suggested in Chapter 4. It is possible that students find the progression through levels, even within cycles, difficult because there are less chances to engage in meaningful activities for data collection than the other focus areas.

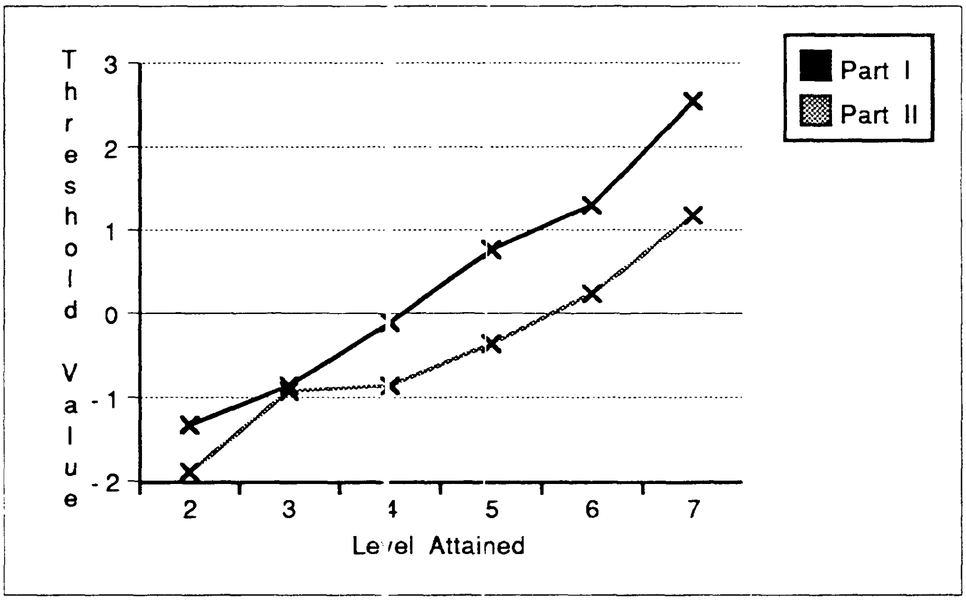

Figure 5.6 - Thresholds to Attain Levels - Question 1

However, the next three graphs, in Figures 5.7 to 5.9, show a very interesting trend. For all three questions the threshold values for Levels 3, 4 and 5 are almost the same. This indicates that once a student can give a Level 3 response it is likely that they will also be able to give a Level 4 or 5 response. These results help to confirm that these three Levels, 3, 4 and 5, are in fact one cycle within a mode. In all three questions the threshold for Level 6, the entry level for the second cycle of the mode, is greater indicating more understanding is required to be able to respond at this Level. The level of understanding required to respond at Level 7, however, is quite high in all questions and mostly quite a bit higher than Level 6. This indicates that secondary students, if they make it to this cycle of the concrete-symbolic mode, need a high level of understanding to proceed through this cycle. In fact the numbers of students responding at Level 7 or 8 were small.

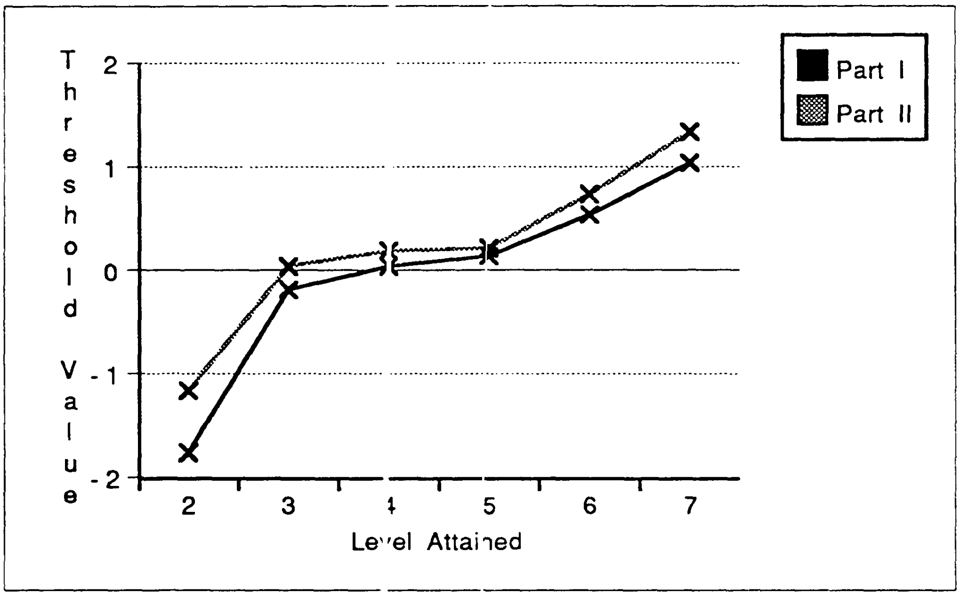

Figure 5.7 - Thresholds to Attain Levels - Question 2

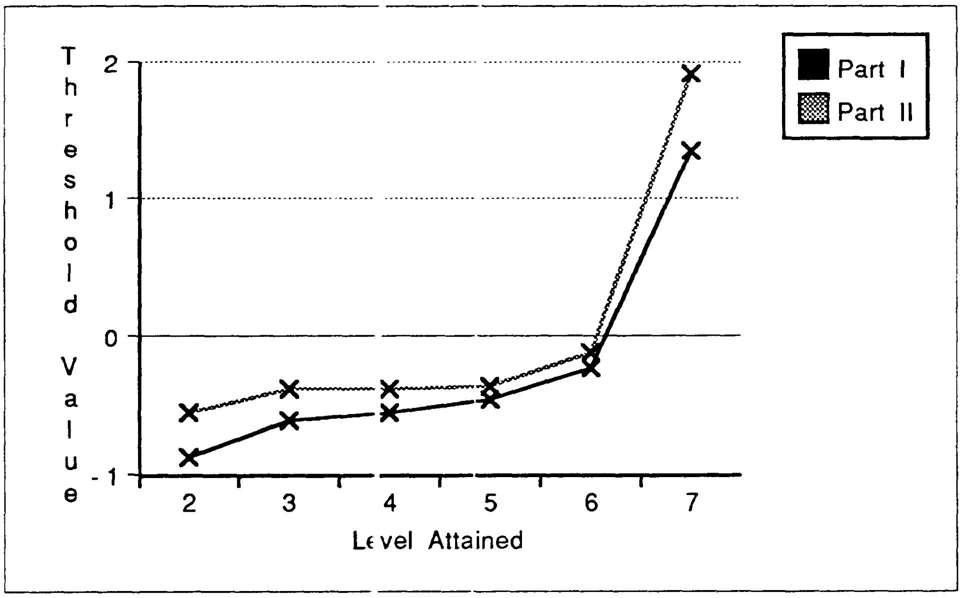

Figure 5.8 - Thresholds to Attain Levels - Question 3

In Figures 5.7 and 5.8, the Part II thresholds are slightly higher indicating that, for the tasks of data tabulation, representation and reduction to respond at a certain level maybe require a little more understanding when the data are presented as a graph rather than in raw form. However, the differences between the values for Part I and Part II are only slight and for Question 4, in Figure 5.9, they are almost identical, indicating that for interpretation and inference the form of presentation of the data does not appear to affect the level of understanding required to present responses at specific levels.

In conclusion, the similarity of threshold values suggests strongly the existence of a cycle at the Levels 3, 4 and 5, at least for three of the focus areas. This is the first cycle within the concrete-symbolic mode. Further reinforcement for the existence of this cycle is given by the earlier tau values which showed the steps into Levels 4 and 5 to be relatively easy compared to the step into Level 3. The second cycle which was suggested in Chapter 4 is not quite so clear. The thresholds for Level 6 are greater than for Level 3, 4 and 5 showing that more understanding is needed to attain this level which is the entry point for the second cycle.

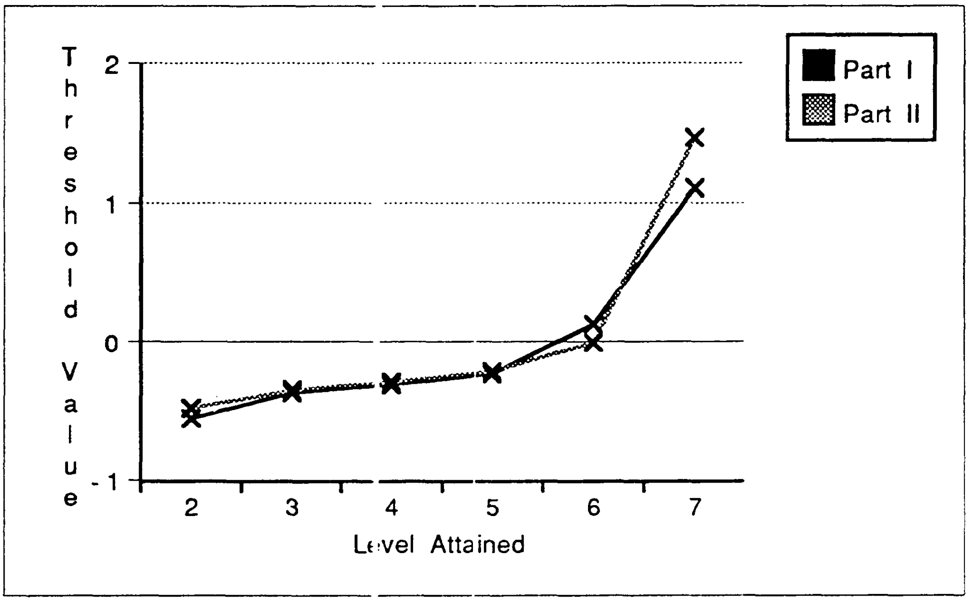

Figure 5.9 - Thresholds to Attain Levels - Question 4

However, Level 7 mostly has an even higher threshold level. These tau values show also that in all questions, except Question 2, the step into Level 7 was even more difficult than the step into Level 6. As there were less responses at this upper end of the hierarchy, further research perhaps including tertiary students, may help to clarify the existence and character of this second cycle in the concrete-symbolic mode.

These threshold values, or estimates for entry into levels, are also useful for determining the stages students may be expected to have attained during their secondary schooling. In the next section, estimate means for students are compared to the threshold levels to provide a profile for the 'average' student.

Profile of an Average Student

Where does this analysis place the *average* secondary student, as far as statistical understanding is concerned? Figure 5.10 was constructed using the level thresholds for the four questions. It is a combination of the information presented in the graphs in Figures 5.6 to 5.9. Each number on the graph represents the level which a student has a 50% chance of entering, if his or her estimate is at the value shown. The position of the 'average' student is shown as the estimate of 0.2 using a - - - - - line. Of most interest are the thresholds for Levels 3 and 6 which are the entry points to the first and the second cycles, respectively, of the concrete-symbolic modes. The diagram shows similar entry estimates for Level 3 except for the data tabulation and representation questions which are a little higher. However, the entry estimates for Level 6 show much more variation. It appears that a lower estimate of understanding is needed to achieve entry into the second cycle in the data reduction and interpretation and inference, than for the data collection and tabulation and representation areas. The line showing the position of an *average* secondary student indicates performance at a higher level in the last two focus areas where responses are already showing features of the second cycle.

In the data collection focus area, the 'average' student is still in the first cycle of the concrete-symbolic mode, attempting a unistructural response of the second cycle, if prompted with some data collection details. In data tabulation and representation, he or she shows good understanding in the first cycle of the concrete-symbolic mode and again is ready to venture into the second cycle. In the data reduction and interpretation and inference focus areas, the average student has just entered the second cycle of the concrete-symbolic mode and is now able to use the data given to justify responses.

However, no student is really *average* and so how does this picture of the average student vary over the three factors, academic year, mathematical ability and gender. To explore these issues it is necessary to first consider the change over academic years. The average Year 7 student is operating in the first cycle of the concrete-symbolic mode in most areas and just on the point of entering the second cycle in data reduction and interpretation and inference. By Year 8 the average student is functioning well in the second cycle in the areas of data reduction, and interpretation and inference, but still in the first cycle in other areas. Year 9 is when students begin to enter the second cycle in data collection with some prompting. For Year 10 and even Year 11 the level of understanding is similar. It is not until Year 12 that the average student begins to enter the second cycle in the focus area of data tabulation and representation. Although differences between academic years are small, as mentioned earlier, there is a big difference between the threshold levels for the average Year 7 and Year 12 students. The overall effect is that, over the six years of schooling, many students move from the first cycle to the second cycle in the concrete-symbolic mode.

To avoid overcrowding on Figure 5.10, the position of the average students for different mathematical abilities and genders were not shown. These are discussed now but it should be remembered both these factors are averaged over academic years and include a wide age range of students as evidenced by the estimates shown in Figures J.5 and J.6.

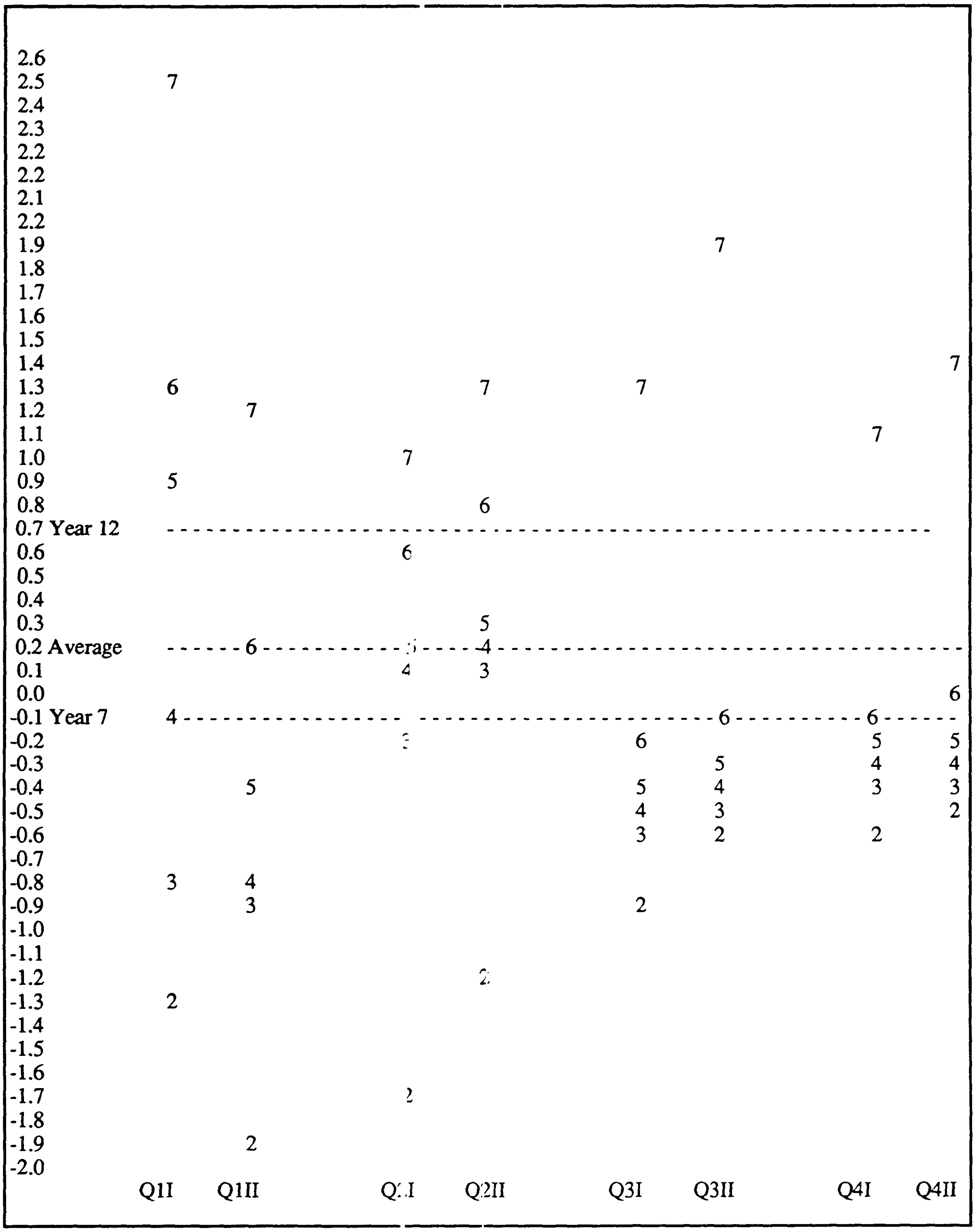

Figure 5.10 - Threshold Value Comparison for Levels within Questions

The average, low mathematical ability student is just above the average Year 7 student, while the average middle and high mathematical ability students are in similar positions just above the 'average' student. For the data collection focus area, students are responding in the first cycle irrespective of their mathematical ability. However, when given some prompting the middle and upper mathematical ability students are beginning to respond in the unistructural level of the second cycle. The lower mathematical ability students are just moving into the first cycle for the data representation and interpretation area while the middle and upper mathematical ability students are operating well in the second cycle. Average students for all mathematical abilities are operating well in the second cycle for the data reduction focus area but this is only true for middle and high mathematical ability students when it comes to interpretation and inference. For this last focus area, the low mathematical ability students are still trying to move into the second cycle.

Thus, as far as mathematical ability is concerned the average, a low mathematical ability student is generally operating in the first cycle, just making the the move into the second cycle in data reduction and interpretation and inference. The average, middle and high mathematical ability students are operating well in the second cycle in these two focus areas but still only just beginning to enter into the second cycle in the data collection area, with prompting.

The average male student is just below the 'average' student while the average female is just above. Hence, as far as the average male and female students are concerned they have similar performances over most of the focus areas. In data collection he or she is operating in the first cycle and moving into the second with prompting. In the data reduction and interpretation and inference areas. he or she is operating comfortably in the second cycle. However, for the data tabulation and representation focus area the average female student is moving into the second cycle while the average male is still responding in the first cycle. The greatest differences between the average male and female are in the entry of the female average only into the second cycle data collection with prompting and the position of females at the upper limit of the first cycle for data representation and tabulation while the average male is still at the beginning of the cycle.

The threshold levels presented, strongly support the suitability of the description of the first cycle in the concrete-symbolic mode, except perhaps in the data collection area. Further research is needed to reinforce the structure of the second cycle. The student profile presented suggests a general trend to move from the first to the second cycle of the concrete-symbolic mode during secondary schooling.

In the next section, the results and conclusions presented in Chapters 4 and 5 are used to directly address the research questions outlined at the end of Chapter 2.

Conclusion

In Chapter 4, levels were defined for the coding of responses in each of the four focus areas. Care was exercised when coding responses as evidence of student understanding is not always found in the part of the question where it was expected. For example, when answering Question 4, some students responded that a pattern could not be found but evidence of identification of a pattern showed through during the justification of predictions, in a later part of the question. The resulting levels attained by students were analyzed to determine the possible influence of the students' academic year, mathematical ability and gender on statistical understanding. Also, frameworks based on the SOLO Taxonomy were suggested for describing this understanding. The present chapter combined the results for the four focus areas and used them to compare question difficulty and threshold levels. These results helped to explain the modes, levels and cycles identified in Chapter 4 and to create a profile of an 'average' student.

The research questions outlined at the end of Chapter 2 are now addressed using the results and conclusions from Study 1. The focus of these questions fell into two main areas of interest, namely: description of a possible hierarchy; and, consideration of factors which affect understanding. Following this, there is a discussion of other issues which arose during the Study 1 research, including: the effect of prompting; the effect of the form of data presentation, which arose out of the trialling of questions in the pilot study; and, an explanation for the different processing paths, which arose out of the analysis of the results of the main study.

The Hierarchy and SOLO Description

The main concerns, initially, were whether there was a hierarchy of growth of students' understanding in statistics and whether SOLO could be used as a framework to explain the hierarchy. These concerns were expressed in questions Q2.1 and Q2.2 at the end of Chapter 2.

In fact, four focus areas were identified within statistics, and hierarchies of growth were described within each of these areas. The feasibility of the design of such hierarchies was indicated early on with the grading of responses during the trialling of questions. Hierarchies identified fell into three main groupings based on the revised framework of statistical thinking as detailed in the results of the pilot study in Chapter 3. Each hierarchy had a possible nine levels spread across the three groups, ranging from a no response at Level 0 to a response at Level 8 which exhibited a reasonably high level of statistical thinking. A number of facts combined to suggest that the hierarchy does describe increasing statistical sophistication. First, the aspects of statistical thinking discussed in Chapter 3 were observed to appear in the responses. Second, the descriptions of the levels showed an increasing awareness and use of data and related properties. Third, and last, there was an overall increase in the level of the responses with increasing academic year as would be anticipated, although the change from year to year was only slight.

Once these hierarchies were identified, it was possible to use the SOLO Taxonomy to explain the various groupings and levels. The basic SOLO framework used was that identified in the Qualitative Data Analysis section for the main study in Chapter 3. This taxonomy was found to be particularly suitable for describing the observed hierarchy because successive levels were identified within the groupings of the hierarchy which deal with individual features, multiple features and then integration of the features as required by the unistructural, multistructural and relational levels, respectively. Both the ikonic and concrete-symbolic modes were recognizable in the groupings of the hierarchy. However, no responses were identified which could be considered to exhibit formal characteristics. Following is a discussion of the modes, levels and cycles identified.

Modes, Levels and Cycles

Responses which were clearly ikonic were identified in all four focus areas and were contained in the first grouping. These responses indicated no concern for the process of data collection or for the use of data presented in the question. Most justifications given were from personal experience and often information given in the wording of the question was reproduced in the answer. Concrete-symbolic responses were identified, also, in all four focus areas. These responses were contained in the second and third groupings. This style of response demonstrated concern for the data and its characteristics and attempts to understand data and use it. There was a considerable range in the success of these attempts.

In the SOLO model, each mode consists of various levels, and the next major concerns identified were the descriptions of the SOLO levels within these modes and the identification of possible cycles within the SOLO modes as expressed in questions Q2.3 and Q2.4. Detailed descriptions of the levels identified within the modes, along with descriptions of particular responses typical of each level, appear at the end of each focus area analysis in Chapter 4. Brief descriptions of the modes, and cycle within them, follow. The ikonic mode was identified in general terms. These responses indicated that either the task had not been addressed or, if addressed, answers given were related to personal experience. To be able to better outline descriptions of the levels within the mode more responses would be needed, probably including some by students at the pre-secondary level.

Two cycles of levels were identified in the concrete-symbolic mode and the unistructural, multistructural and relational levels identified within each. The first cycle related to concern for: physical aspects of data collection; basic descriptions of data; reduction of data to descriptive statistics; and, informal predictions not based on data trends. The second cycle related to: appreciation that the method of data selection influences the quality of data; description of data by referring to specific features of the data; reduction of data to formal statistics; and, recognition of patterns in data allowing predictions to be made.

Justification of the splits into levels and cycles are presented in Chapter 5. The discussion of tau and threshold values reinforced the notion that the Levels 3, 4 and 5 were in fact three levels within one cycle and Levels 6, 7 and 8 were within another. The tau values showed that generally the step into Levels 3 and 6, the entry to the first and second cycles of the concrete-symbolic mode, were difficult compared to other levels. The step into Level 7 was also difficult compared to Level 6 and this needs to be studied further with a larger sample of responses. Threshold value analysis showed that Levels 3, 4 and 5 for most of the questions, have similar estimates for entry by students suggesting that these form one cycle. As Levels 7 and 8 were combined and threshold levels were higher than for 6 there is not sufficient evidence to decide as conclusively that Levels 6, 7 and 8 are in one cycle.

A possible explanation for the higher than expected threshold for Level 7 is that this is the point at which many secondary students have reached their upper limit of functioning. Biggs and Collis (1982, pp.22-23) discussed the hypothetical cognitive structure (HCS) which is the generalized cognitive structure of an individual. The HCS is not directly measurable and is different to the SOLO level which exists in the student's response. The levels recognized by this study in the second cycle of the concrete-symbolic mode in SOLO would have more respresentative threshold values if the sample of students had not included only secondary students, most of whom would have reached their HCS upper limit of functioning at that stage. Analysis of a sample of responses including tertiary students, who would have a higher upper limit of functioning, could help to confirm more tightly the levels and structure of the second cycle in the concrete-symbolic mode.

Ikonic to Concrete-Symbolic Transition

Most students at some stage in their primary-secondary schooling make the transition from the ikonic to the concrete-symbolic mode in the various aspects of statistical understanding. Hence, it is important to clearly mark the distinction between these two modes. Question Q2.5 expressed this need. When thinking is in the ikonic mode: data are seen as an endpoint and not something to be used; any justifications given are based on personal experiences and beliefs; and, data collection is not taken into consideration when addressing queries, with data being considered as something that is already readily available. The transition into the concrete-symbolic mode has been made once there is some indication of reaction to the data. The data are now viewed as the beginning of a process and used to lead to something else. This is the stage where there is also the realization that data collection is a process over which control may be exerted.

Factors Affecting Understanding

The possibilities of a student's academic year, mathematical ability or gender influencing level of understanding were proposed in research questions Q2.6, Q2.7 and Q2.8 respectively. The QUEST analysis showed that all three factors affect a student's overall level of statistical understanding. The Scheffe test of pairs of means for these factors and the chi-squared analyses for each of the focus areas allowed detailed discussion of these differences.

Academic Year

Although there is a slight increase in the level of understanding from one academic year to the next, the difference does not become significant until Year 7 is compared to the senior years. Each focus area showed a slight increase in the main 'bulge' of the level of responses from Years 7 to 12, with the least evidence of a shift in the area of data reduction where most academic years had a predominance of responses at Level 6.

Mathematical Ability

The significant mathematical ability factor was shown to be mainly due to a difference in understanding between the low mathematical ability students and the others. In the data collection, and interpretation and inference areas, low mathematical ability students showed less understanding than the middle and high mathematical ability students. However, in the areas of data collection, and data tabulation and representation, there was only evidence of poorer performance by these students in the prompted data collection situation.

Gender

The significant gender factor showed up as a higher level of statistical understanding for females than males. However, when individual focus areas were analyzed, many of the results showed similar distributions of levels of responses for males and females. The female students showed a higher level of understanding in the prompted data collection situation and in the data reduction question when the data were presented as a graph. The results for the interpretation and inference question, with the data presented in raw form, were significant also, but only just. Hence, no strong conclusions should be made under the circumstances.

In response to the research questions, all three factors have been found to influence the level of statistical understanding, with the influence of academic year and mathematical ability perhaps being more significant, and of more consequence, than the influence of gender.

Other Considerations

As mentioned earlier, as well as the main research questions, other queries arose during the course of the study. Trialling of questions suggested that the effects of prompting, and the form of presentation of the data, may influence the level of understanding exhibited by students' responses. These issues are addressed below, along with the question of the existence of different processing paths which became apparent during the analysis of the responses in Study 1.

Prompting

The question of the effect of prompting arose out of the trialling of data collection questions. The data collection question decided on for the final test was designed so that the first part had a general data collection question and the second had a situation where some prompting was given as to some of the practical aspects of the data collection to give students a chance to concentrate on finer details. Comparison of the results of the two parts of this question showed that the prompting of some information concerning the data collection process increased significantly the level of the response. This effect became more noticeable in the later academic years. Also, females and students with high mathematical ability gave higher level responses when assisted with a prompt. The effect of the form of presentation of data was not as significant as the effect of prompting but it did influence the level of understanding shown in the responses and this is discussed next.

Form of Data Presentation

The question of the possible effect on the level of understanding of the presentation of data in different forms also arose out of the trialling of questions. For the written test, all questions which dealt with data were designed with two parts which presented the data first in raw and then in graph form. Comparison of the results of the two parts of the questions showed a lower level of understanding in the data tabulation and representation, and data reduction areas, when the information was presented as a graph. However, for the interpretation and inference task similar levels of understanding were exhibited irrespective of the form of data presentation.

The academic year and mathematical ability of the student did not appear to influence these results. Gender, on the other hand, did in some areas. Females showed a higher level of understanding than males when data were presented as a graph for the data reduction task and when the data were presented in raw form for the interpretation and inference task. As well as these issues which arose during the trialling of questions, the analysis of responses showed that there were parallel processing paths which could be taken within each level. Once these paths were recognized to exist the analysis was adjusted to incorporate a component for testing the differences between students' choice of path.

Processing Paths

As responses were coded into a hierarchy of levels for each question that dealt with data, it became apparent that the more statistically sophisticated responses tended to diverge into two parallel processing paths. For the data tabulation and representation question, these two paths involved using either statistics or judgements as the basis for the description of the data. Similar numbers of students chose to answer the question using each processing path when the data were presented in a table. With graphical presentation of data, more students chose to use statistics as the basis for the descriptions. These results were not influenced by either the gender or mathematical ability of the student. The only significant influence of mathematical ability was that, when data were presented graphically, the lower mathematical ability students were more likely to use judgements rather than statistics in their descriptions.

For data reduction and interpretation and inference, the two processing paths relied on either measures of central tendency or measures of dispersion. Far more central tendency-based responses were recorded in the data reduction question, while far more dispersion-based responses were recorded in the interpretation and inference question. The gender or mathematical ability of the student did not appear to affect the choice of processing path in either the data reduction or interpretation and inference questions. With higher level responses, the two paths converge again and students attempt to use one type of measure to qualify the discussion of the other. Research into the link between the mean and standard deviation was mentioned in Chapter 2.

Most importantly, students tended to follow the same processing path for both tabular and graphical presentation of data, except in the interpretation and inference area where there appears to be no relationship between the processing path chosen and the form of data presentation.

Having addressed the issues on which Study 1 was based, as well as some which arose during the course of the research, the remainder of the research questions posed at the end of Chapter 2 are addressed in Study 2. The second study was intended as longitudinal research. However, extra dimensions investigating prompting and processing paths, were added to the study as a result of the analysis undertaken in Study 1.

CHAPTER 6

DESIGN - STUDY 2

Some of the research questions presented in Chapter 2 were intended to be addressed in the second study. The main focii of these were the robustness of the levels of the hierarchy, the growth in understanding over time, and the possible elaboration provided by oral testing. In addition, the divergence of responses into parallel paths, observed in the first study, warranted explanative investigation. In particular, the existence of a link between the path of the response and the processing style, successive or simultaneous, of the student was considered. To answer these questions students were retested 12 months after Study 1, and some were interviewed.

Study 2 consisted of three main parts, a Luria Study, a Longitudinal Study and an Interview Study. The object of this chapter, which contains four sections, is to describe the development of the experimental design of the second study and the methodological considerations incorporated into the study. The first section, **Background**, outlines the context of Study 2 as well as the limitations and constraints imposed. It also includes a short discussion of successive and simultaneous processing and Luria style testing. The second, **Luria Study**, describes how students were tested for processing style in an attempt to explain the divergence into separate processing paths observed in the first study. The next, **Longitudinal Study**, describes the process used to retest all available students one year after the first study, including the sampling methodology and data analysis plan. The concluding section, **Interview Study**, outlines the process used to interview a small selection of students to expand on the information provided by the written test. The results of all components of Study 2 are presented in Chapter 7.

Background

As a consequence of diverging processing paths in the responses in Study 1, Study 2 was expanded from its original format to include an investigation of simultaneous and successive processing. This section presents a short discussion of simultaneous and successive processing and the Luria style tests used to measure this characteristic. Following that, the context of Study 2, the limitations and constraints imposed are outlined.

Successive and Simultaneous Processing

Both Biggs and Collis (1982, p.193) and Collis (1986, p.10) suggested that learning within modes can be related to simultaneous versus successive processing. This prompted the investigation into these styles of processing as possible explanations for the branching of the responses in Study 1. Various researchers have demonstrated that where learning tasks have been related to simultaneous versus successive processing, differences in learning outcomes have been observed (Angus, 1984, p.6). Luria (1966) has extensively researched the functioning of the brain and maintains that there is strong evidence to suggest there are "two basic forms of integrative activity by the cerebral vertex" (p.101). These two forms are called simultaneous and successive. In the first, individual stimuli arriving in the brain are integrated into simultaneous groups while in the second, the stimuli are arranged into a successive series as they consecutively arrive.

Angus (1984) discussed these two styles of processing and used Luria's model and battery of tests as a basis for his research into designing tasks and investigating their performance. The Luria tests on which Angus based his research were used in the Luria Study section of Study 2 to investigate processing style. The key research question was whether there was a link between a student's tendency to process simultaneously or successively and their tendency to provide a response which demonstrated processing in the A or B path as distinguished in Study 1. This investigation was incorporated into the research framework originally planned for Study 2.

Context

The same school population was used for Study 2 as for Study 1. Aspects of this population are outlined in Chapter 3, and include such information as the place of the research school in the education system and the exposure of the students to statistics. The main change in context from Study 1 is that the students are one year older. Luria Tests were applied early in the year. However, by the time that the written test for the Longitudinal Study was applied, a full 12 months had elapsed since Study 1. The interviews took place as soon as possible after the written test. Permission for further testing was sought from parents, a copy of the letter to parents is reproduced in Appendix K.

Limitations and Constraints

The limitations and constraints imposed on Study 1, as described in Chapter 3, also applied to Study 2. They included influences on both the sample, such as students' choice to be involved, and on the procedure, such as the distractions involved with testing in a normal classroom situation. As far as the sample was concerned, students were again given the option of whether they wished to be part of the study, and some students previously involved chose not to be included in Study 2. Normal classrooms were used, as with Study 1, so that students would feel more comfortable about doing the test. All testing was done during school hours and at times which were mutually suitable for the students and the researcher. These limitations and constraints, however, were not considered detrimental to the study. Next, the details of the first component of Study 2, the Luria Study, are presented.

Luria Study

As mentioned earlier, Study 1 uncovered a trend in a number of the levels of the response hierarchy where the detail in the answers split into two seemingly parallel paths of processing, labelled A and B. The Luria component of Study 2 was designed to determine whether the differing paths of responses, observed in Study 1, were related to the fact that some students process successively while others process simultaneously. First, it was necessary to determine a student's tendency towards simultaneous or successive processing. Then, the possibility of a relationship between this tendency and the tendency to process along the A or B paths was investigated. This section details the participants and instrumentation for this component of Study 2, and the data analysis plan.

Participants and Instrumentation

The Luria Study aimed to apply the Luria Test to as many students as possible from the original sample of students who participated in Study 1. The sample selected, the test instrument used and the procedure involved with applying the test instrument are detailed in this section.

Sample

Of the original 180 students, 97 were tested in the Luria Study. Not all students were available for this phase of the study as some had left school, including the 30 Year 12 students who had graduated. The final proportion of females to males was considered to be appropriate although there was a slight imbalance in Year 8. Unfortunately, some students chose not to participate in the testing, the response from the new Year 12 being particularly disappointing. This meant that senior students were under-represented. However, as the proposed investigation was not expected to be more relevant in any particular year this problem with the sample was not considered to detract from the overall research.

The year given is the academic year of the student in the second year of the study, for example a Year 7 student for Study 1 is now a Year 8 student. Hence there are no

Year 7s, as they were not in the school the year before. The breakdown of this sample by academic year, mathematical ability and gender, is shown in Table 6.1.

Table 6.1 - Luria Study Sample

	Mathematical Ability						
	Low		Middle		High		Total
Year	Female	Male	Female	Male	Female	Male	
8	5	2	2	2	5	3	**19**
9	4	3	4	5	4	5	**25**
10	4	5	3	3	3	5	**23**
11	3	3	4	4	3	3	**20**
12	2	2	3	3	0	0	**10**
Total	**18**	**15**	**16**	**17**	**15**	**16**	**97**

Research Instrument

A combined written and verbal test was administered to the students, consisting of six separate sets of test items. These were selected from a larger range of possible Luria style test items, as used by Angus (1984), and covered a range of tasks which required either successive or simultaneous processing to maximize the number of successful responses. All answers were recorded in written form but some questions were asked verbally and others were presented in printed form.

The selected tests are reproduced in Appendix L and the title and time limit for each are shown in Table 6.2. Tests numbered 3, 4 and 6 are aimed at measuring successive processing capabilities while tests numbered 2 and 5 measure the simultaneous factor. Test numbered 1 is mainly aimed at measuring attention span but does have some ability to discriminate successive processing tendencies (Angus, 1984, pp.78-81).

Before each of the six tests, students attempted a sample item, with a full explanation of how it should be completed, and they were given the chance to ask any questions. These sample items, together with the administration instructions, are included in Appendix L with the test pages that were distributed to the students. Tests numbered 1, 2 and 5 required students to complete written test pages in the specified time. Tests numbered 3, 4 and 6 were administered verbally, with students recording a written answer to each item. This meant that students were given a time limit per item rather than a time limit per page. The time limits given in Table 6.2 were the overall time limits for an entire page.

Table 6.2 - Luria Test Bank

1	Number/Letter Search Test	3.5 mins
2	Form Board	8.0 mins
3	Auditory Word Span	7.0 mins
4	Number/Letter Attention Span	5.0 mins
5	Paper Folding	2.5 mins
6	Auditory Number Span	6.0 mins

The test was administered during normal lesson time in a standard classroom. Students were tested in groups comprising one academic year at a time. As some questions were to be administered verbally, a classroom with as little outside distraction as possible was selected. A short break was left between each test, as relief from the intense concentration students needed to apply under the pressure of the time constraints. Students were allowed to ask questions after the completion of the test item for each page. Students were not allowed to go back to a previous page once the time limit was up. The responses to the Luria Tests, along with the responses to the written test in Study 1, were analyzed to investigate the relationship between processing style and response path.

Data Analysis Plan

There were two distinct stages in the data analysis, namely: the processing of the Luria test results; and, then the comparing of these results to those of the written test from Study 1. The processing of the Luria results was needed to determine whether the Luria scores could discriminate between those students who processed successively and those who processed simultaneously, and, if so, the strength of the tendency to process using each of these styles. The students' responses to the test items were marked for correctness and then their scores for each of the six test pages were used to produce two coefficients. One to measure the degree of successive processing the student had exhibited, and the other to measure simultaneous processing. A factor analysis was used as a check on the reliability of the test as a measure of these variables.

Before comparing the paths chosen and the Luria scores, the students' levels of understanding were analyzed with the Luria factors to determine whether the style of processing influenced the level of understanding exhibited in responses. A multiple regression of the Rasch score on the successive and simultaneous scores was used.

For the second stage, the comparing of the path chosen to the Luria scores, a measure was needed of a student's tendency to respond along Path A or B. The questions which involved working with data, Questions 2, 3 and 4, were those where the divergence was observed, so the responses to Question 1 were not considered when determining a Path score. Each of the three questions had two parts so there were six possible answers for which a student may have given an A or B path response. However, low level responses (Levels 0, 1, 2, 3 and 4) and high level responses (Level 8 and some Level 7) did not exhibit the split. Hence, all students had an indication of preferred path for every one of the six responses. The Path score was designed as a measure of the preference for Path A over Path B. This score was calculated by subtracting the number of B responses from the number of A responses and dividing by the difference between 7 and the total number of answers which had a path response of some type for that student. This gave a score from -6, indicating a Path B response to all six, to +6, indicating all Path A responses. The division factor was introduced so that a person who gave answers which were only from one path obtained a higher score than someone who gave answers from both paths. For example, using a simple subtraction as the score, a person with one A path response would receive the same score, 1, as a person who answers with three A paths and two B path but dividing by the factor gives the first person, who showed no tendency to the B path a higher score, 1, than the other who would receive 0.2. This Path score was then analyzed with respect to the Luria scores. First, a multiple regression was performed as an indication of the overall effect of the two factors. Then, each factor, successive and simultaneous, was separately regressed against the Path score.

Finally, an indication of the association between the path chosen and the successive and simultaneous processing variables was calculated, separately, for each of the questions. Due to the reduced number of pertinent data items rather than use the processed scores for the successive and simultaneous factors, the students were labelled as low, middle or high with respect to their tendency to process successively or simultaneously. A score ranking from 1 to 3 was used to replace each of the coefficients. The 40% of students with the lowest scores received a 1, the higher 40% a 3 and the middle 20% received a 2. Using these scores students with a 1 were considered to be Low in whichever processing style was being represented and a 3 was considered to be High. For each question, these results were tabulated against the A or B path for all the students with suitable responses. A Chi-squared Test of the Contingency Table for each question was used to determine whether there was an association between the style of processing and the tendency to give an A or B response.

The results of these analyses are presented in the first section of Chapter 7. The next section describes the participants and instrumentation, and data analysis plan for the Longitudinal Study.

Longitudinal Study

The Longitudinal Study was planned to investigate the stability of the hierarchy, that is, verify the results of Study 1, and analyze students' growth in understanding over time. In all, there were three main issues, namely: to check whether the groupings identified in the first study were sufficient to categorize the variety of possible responses; to determine the robustness of the hierarchy developed in the first study as alluded to in Q2.9 in Chapter 2, that is whether the performance of students was consistent over time; and, to investigate whether the level of understanding had increased over the twelve-month period. This section begins with the details of sampling, followed by the data analysis plan.

Participants and Instrumentation

The aim of the study was to apply a four-question written test to the same group of students as tested in Study 1. Details of the sample selected, the test instrument used and the procedure for its application are now described.

Sample

As many students as possible were selected from the 180 students who had originally participated in Study 1. As Year 12 had graduated, there were at most 150 students available, 30 from each of the new Years 8 to 12. However, of these, some had left the school and some did not wish to participate further. The stratification of the sample is presented in Table 6.3. This sample was considered to be sufficiently representative, as the number in each category of the partition of the sample never went below two survivors out of the original five selected for Study 1.

Research Instrument

The written test used the same four questions as for Study 1, with slight alterations. For example, if a question referred to Year 7 before, the new version concerned Year 8, or if the question in Study 1 dealt with data collection of the number of icecreams that had been eaten before, the new version dealt with the number of times a bike had been ridden. This was done to ensure enough similarity in the questions to allow comparison but enough dissimilarity to prevent a sense of déjà vu for those who remembered the test from the previous year. The written test for the Longitudinal Study is reproduced in Appendix M.

Table 6.3 - Longitudinal Study Sample

	Mathematical Ability						
	Low		Middle		High		Total
Year	Female	Male	Female	Male	Female	Male	
8	5	3	2	3	5	3	21
9	4	4	5	5	4	5	27
10	4	4	5	2	3	5	23
11	3	3	5	4	5	3	23
12	4	3	4	4	4	4	23
Total	**20**	**17**	**21**	**18**	**21**	**20**	**117**

As with Study 1, the written test was administered to each year group in a classroom situation. Similar instructions and encouragement were given. Students were also asked to include in their responses as much information as possible.

Data Analysis Plan

The analysis of data for the Longitudinal Study included validating the test instrument, verifying the results of Study 1, and comparing individual student results to consider growth in understanding. When comparisons were to be made with the Study 1 results, from 1991, only those students were used who participated in the retesting in 1992. When this group was used it will be referred to as the Reduced 1991 sample and new estimates of understanding used are produced using only the responses of the reduced sample. When making comparisons, the 1991 and 1992 results were considered independent, as enough time, one year, had elapsed for the 1992 responses not to be influenced by the 1991 responses. The hierarchy developed in Study 1 was used to code the responses and the resulting data were tabulated and tested. These were also processed using the QUEST system to produce a measure of statistical understanding and the necessary statistics for testing the validation of the test instrument.

Validation of Test Instrument

As results were to be compared to those for Study 1, it was necessary to check whether the test was a useful instrument to measure understanding and whether the measurement was similar in performance to that of the Study 1 written test. The same three measures, as for study 1, were used to determine whether the test was a useful measure, and fitted the Rasch model, and whether each question made significant a contribution to the model. These were the item consistency, the parameter fit statistics for each question and student estimates, and the infit mean square maps.

To determine whether the slightly modified test, as used in 1992, was comparable in performance to the 1991 test, tau and threshold values, as produced by the Rasch analysis, were compared. The tau value measures the level of difficulty experienced by a student in responding at a particular level for any question and the threshold value measures the minimum level of understanding a student must have reached to have a 50% chance of responding at a particular level. Similar analysis was performed for each of the two comparisons. Results for the 1992 sample were compared to results for the Reduced 1991 sample, rather than the entire 1991 sample, so that the same students were used for each year to reduce the risk of any differences being due to variation within the sample of students. To test whether there was a relationship between the two years, the Pearson correlation r was calculated for all levels combined over all questions and tested for significance using the Student's t distribution for $t = r \div \sqrt{(1-r^2) \div (n-2)}$ on $(n-2)$ degrees of freedom. To further investigate the relationship, the correlations were analyzed separately for each of the four questions, again calculating the correlation r and testing it for significance.

Verification of Study 1 Results

The 1992 data were analyzed in a similar fashion to the 1991 data, including subjective discussion of the influence on the levels and paths of the academic year factor and chi-squared analyses of the influence of the mathematical ability and gender factors. Also included were the chi-squared comparisons of the two parts of each question and the analysis of variance of the Rasch scores measuring overall statistical understanding.

Once these analyses were completed, the level of significance of the tests were compared to those observed in Study 1 to determine whether the results and conclusions for Study 2 confirmed those for Study 1. The plan for analysis of the change in understanding of individual students over time is presented next.

Longitudinal Comparison

To check whether there were major reversals in performance the correlation of 1991 and 1992 estimates was calculated. All other analyses relevant to the growth in understanding over time were performed on the differences in understanding for students. These were calculated by subtracting the 1991 estimate from the 1992 estimate, so a positive difference indicated an increase in understanding. T-tests were used to determine whether the differences were significant, firstly for the entire sample and then separately for each academic year, level of mathematical ability, and gender. An analysis of variance was also applied to the differences to determine whether the factors academic year, mathematical ability, and gender affected the change in overall statistical understanding over time. The results of the analysis are presented in Chapter 7.

Interview Study

This study consisted of a detailed interview session with a representative number of the students who had been part of the Longitudinal Study and was designed to address two main research questions. First, it was necessary to determine any differences in responses which may occur due to answering orally. This query was alluded to in research question Q2.10 in Chapter 2. Did the student give a response which was optimal or was it possible that the student had not made a complete disclosure? Second, the effect of prompting was investigated in the interview situation. Reference to this research area was made in question Q2.11. More specifically, the interviews were designed to determine whether the students were able to present responses which demonstrated a greater understanding than that initially coded, either under probing or prompting?

The interviews provided a means of checking whether more detail was available from students under controlled probing and prompting. Probing, involves encouraging a student to add to a response, without giving any assistance towards answering the question. However, prompting is encouragement which includes some qualification or information which assists in answering the question.

Participants and Instrumentation

The study was designed to interview a small selection of students who had participated in the Longitudinal Study. The sample selected and the research instrument used are outlined in this section.

Sample

As it was not possible to interview all of the students who completed the written test, 15 students were purposively selected. From each of the Years 8 to 12, 3 students were selected randomly from the sample tested in the Longitudinal Study so that overall the ratio of females to males was as close as possible to 1:1 and in any one year there was at least one student of each sex. The resulting sample breakdown was 7 female and 8 male students. In terms of the mathematical ability levels, low, middle and high, there were 2, 7 and 6 students, respectively.

Research Instrument

The interviews were controlled work-throughs of the test already completed. The schedule devised for the interview, which appears in Appendix N, was based on the Newman strategy for interviewing (Clements, 1980, p.9). This sought to determine reading, transforming, encoding and careless errors. Also, the students were probed to determine whether their first attempt at a response contained all information that they were able to provide. Then they were prompted to determine whether the level of the response could be raised with some assistance.

During the interview the complete written test was worked through, question by question, with the same procedure being followed for each question. Firstly, reading and comprehension and interpretation phases were checked. The student was asked to read the question and then asked to explain what the question meant. Next, came a discussion of the strategy and process selected by the student. The student was asked for a response to the question, explaining his or her answer as well as possible. If there was a major misunderstanding in the interpretation of the question, it was dealt with at this stage to allow the student to present his or her best possible answer. The student was given a reasonable length of time to begin a response, about 15 seconds, this time period being similar to that used by Allwood and Montgomery (1982, p.132).

The student was then probed to see whether more could be given in the answer. Some probes used were Good....?, Anymore....? and Can you add anything to that.....? Again, a student was not left silent for more than 15 seconds, more probing being used to encourage any addition to the response. If the response was different to that given in the written test, the previous answer was read and he or she was asked to choose between the two answers, and give reasons for the choice. Finally, the student was prompted to see whether any more depth could be added to the answer by having various ideas suggested to them or problems clarified.

The interview schedule was used as a guide, and questions which needed to be asked varied slightly depending on each individual subject's characteristics and earlier responses. Contingency, the expectation that two researchers would get exactly the same answers interviewing the same student was not considered to be a problem in this situation. Discovery of the possible existence of a higher level of understanding than exhibited in the written test was the aim of the interviews rather than a strict measure of variables, making the issues of the reliability and validity of the measurement instrument not relevant. This is a similar situation to that of the discovery type clinical interviews conducted by Ginsburg, Kossan, Schwartz and Swanson (1983).

The interviews were conducted on a one-on-one basis in a private area. This was done within a few weeks of the student completing the written test, so that the test was still fresh in the student's mind. The session was tape recorded with students initially being given the chance to hear their own voice on tape to reduce tension.

Data Analysis Plan

Analysis of the interviews was both quantitative and qualitative. Transcripts of each interview were prepared and from these each student was allocated three separate coded levels for each part of each question. These were, an initial interview response level, a level achieved after probing and a level achieved after prompting. The changes in level between written test and interview and the changes in level due to probing were tabulated to investigate any differences between responses due to oral testing rather than written. The changes in level between the probed response and prompted response were tabulated to demonstrate any improvement in level due to prompting. Separate tallies for the factors academic year, mathematical ability and gender are not presented as the sample size of only fifteen makes cell sizes in such tables too small to be statistically useful.

Qualitative assessments were made of any problems with reading, comprehension and interpretation of questions, and also any interesting aspects or anomalies with respect to the measurement of level of understanding which arose during the interviews. Due to the number of students interviewed, and the nature of the research questions asked, no formal statistical tests were applied to the response level changes resulting from the coding. The only quantitative data analysis was the creation of tables of response levels as described above.

Conclusion

The three studies, Luria, Longitudinal and Interview, combined to form Study 2, and were created to reinforce and extend Study 1. The design of Study 2 was planned so that the sample was the largest possible subset of the Study 1 sample. The analysis of results was linked to the analysis of Study 1 results and used coding hierarchies created in Study 1. The results of the three sections of Study 2 are presented in the next chapter and then results and conclusions from both Study 1 and Study 2 are discussed in Chapter 8.