An exploration of students' statistical thinking using the SOLO Taxonomy.

Chris Reading and John Pegg
University of New England

A paper presented to the 25th Australian Association for Research Conference in Education, November 1995 in Hobart

Abstract

There has been a growing recognition throughout Australia of the place of statistics in the school curriculum. Part of the impetus for this has been the strand status offered statistics (within Chance and Data) in A National Statement in Mathematics for Australian Schools. As expected, this change of emphasis in the curriculum has been mirrored by similar changes in the research agenda. An important aspect of research has been the consideration of what is meant by 'statistical thinking'. This paper takes up this theme by considering students' responses to two open-ended tasks which require the application of data reduction. To assist in this process the SOLO Taxonomy is employed as the theoretical framework.

Introduction

There is a growing interest in Australia and world-wide into the place of statistics in the school curriculum. In Australia, both state and federal educational authorities have indicated that statistical ideas should be incorporated into modern mathematics syllabuses and this should occur across the primary and secondary school years. This important change, and the need to decide where topics should occur, has highlighted the poor research base that exists to guide curriculum issues in statistics education.

In an attempt to address this concern some researchers have begun to explore students' thinking about various aspects of statistics using a neo-Piagetian framework referred to as the SOLO Taxonomy (Biggs & Collis, 1982). This work explores students' responses to selected questions and classifies them in terms of the descriptions associated with the model. Several benefits are possible using this approach, namely,

- qualitative aspects of understanding are explored rather than the usual quantitative ones;
- a broad sense of hierarchical growth is identified;
- links can be framed with other topics in related and non-related areas;

- population trends can be identified.

This paper contributes to this trend by exploring students' responses to statistical questions and considering these in the light of the SOLO Taxonomy. However, before describing the research undertaken it is necessary to provide information on the two contexts that underpin the study, namely, the statistical basis and the SOLO Taxonomy.

Background

A useful description of statistics, for educational purposes, has been developed by Holmes (1980) in his role as director of the Schools Council Project on Statistical Education in the United Kingdom. His

description includes five broad areas which are sub-divided into several sections. These are:

A.Data CollectionB.Data tabulation and representation
use of a censustabulation
types of datapictorial representation
sampling
obtaining data

C.Data reductionD.Probability
measures of locationassigning probabilities
measures of dispersionmanipulating probabilities
other summarising statisticsmodelling and probability distributions
regression and correlation

E.Interpretation and inference
reading information
more detailed interpretation
designing experiments
inference

This list provides a comprehensive summary of the topics of school statistics at which research can be directed.

The SOLO Taxonomy has been described in detail elsewhere (see for example, Biggs & Collis, 1991; Pegg, 1992). In brief, the model comprises two aspects, modes of functioning and levels of achievement, which allow students' responses to be categorised. There are five modes and these represent a growth in abstraction: from reacting to the world by physical actions (sensori-motor); to using imaging and imagination (ikonic); to operating with second-order symbol systems such as written language (concrete symbolic); to being able to deduce general principles and work deductively (formal); to, finally, being able to challenge known theories (post formal). While these modes have much in common with those suggested by Piaget there are differences.

Two are of relevance. The first concerns the placement of Piaget's "early formal" stage into the cycle of levels in the concrete symbolic mode. The second is that the earlier modes are not seen to replace subsequent modes. Instead, earlier modes continue to evolve in their own right and to support growth in later modes. While uni-modal behaviour is possible, quite often, peoples' responses are a complex interaction between a number of modes.

Associated within each mode are a series of levels. Three levels are relevant to the work reported here. They are referred to as

unistructural-a focus on one aspect
multistructural-a focus on several aspects which are unrelated
relational-a focus on several aspects in which inter-relationships are identified

These three levels form a cycle of growth which reoccurs both within a mode and in different modes as a student responds with greater sophistication. In this case of within mode growth, the relational responses of the previous cycle are similar but not as concise as the unistructural level response of the next cycle. When different modes are explored the same pattern of cycles occur although the nature of the element upon which the level is based is obviously different. The value of the SOLO Taxonomy lies in the depth of analysis it provides for interpreting students' responses. The research reported below adopts this framework.

Research Design

This study represents a small part of a much more detailed investigation. The main study involved, in its first phase, an exploration of student understanding of the topic areas A, B, C and E described by Holmes. One hundred and eighty students, 30 from each of Years 7 to 12, were tested on a range of statistical questions. The students were selected randomly from each of the top, middle and bottom third of the population based on their mathematics ability. Within each year group there was a male/female balance.

This paper reports on the students' responses to two questions concerning data reduction. In particular, the area examined included: measures of central tendency including mean, mode and median; and, measures of dispersion including range, quartiles, interquartile range, mean deviation, variance, standard deviation and dispersion.

The questions in the test, provided below, were designed to help clarify students' understanding of the reduction of data into a statistical form. This is one of the earlier steps necessary in the process of analysing data. The aim of the question was to present students with some data and then have them reduce that data into a more

useable form. The questions were left open with no reference to any specific statistics in order to view what students perceived as necessary steps in data reduction and to allow them to use whatever facilities they had available and felt were suitable for the task. There were two forms in which the data was presented. In Part I the data was presented as raw data and in Part II the data was presented as a graph. This was done in order to investigate whether the form in which the data was presented influenced the way in which students reduced the data.

The responses to each part of the question are analyzed separately below before being brought together for an overall analysis of performance. First, there is a discussion of possible levels into which the responses to Part I may be categorized and student performance over the levels. Then the results for Part II of the question are presented. Next, Parts I and II are compared to consider similarities and differences in the student performances and the implications of these observations. Finally, the SOLO Taxonomy is used to suggest a possible framework for explaining the responses and features of the various levels within the modes.

Analysis of Responses to Part I

This open-ended question was designed so that no reference was made specifically to any measure of central tendency or dispersion and students were free to make use of any measure that they felt appropriate. The question, as it was presented to the students, is given in Figure 1 below.

Upon investigating students' responses it was possible to divide them into a number of levels based on the statistical quality of the answer they gave. These levels were able to be grouped together further based on the depth to which the response indicated the ability of the student to cope with the question. Sample responses for the various levels within these groups are presented below.

Question Part I
As part of a large project which had to do with measuring and discussing the human body, one of the tasks was to measure the lengths of peoples' feet to the nearest centimetre. The results of the 29

students in the class are as follows :

26	26	26	27	27	27	27	28	28	28
28	28	28	29	29	29	29	29	30	30
30	30	30	30	30	31	32	32	33	

(i) If you were asked to give a number, or numbers, (to the nearest cm.) which could be best used to represent the size of the left feet

in that class, what numbers, or numbers, would you select ?
(ii) Give reasons for your selection.

Figure 1

First Group

First are those responses which dealt only with the requirements of the question. There were three Levels, coded as 0,1 and 2, observed. These responses indicate consideration of the requirements of the question with no use taken of the data in formulating the response. The following is a brief description of the levels within this group and sample responses from each. The responses have been recorded with the answer to part (i) indicated as such and the answer to part (ii) similarly, (when given by the student).

Level 0These responses indicate that the requirements of the question were not understood or an answer could not be attempted. For example:

(7101)(i)I don't know.
(9209)(i) I haven't got the faintest idea. I haven't learnt it.

Level 1These responses indicate an attempt at answering the question but either it is a nonsense answer or a reason is given which does not answer the question asked. For example:

(7102)(i)I would select all of the odd numbers because the left foot is the same as the left hand, most people are right handed.
(7214)(i)About 4 meters(sic).
(ii)don't know
(8201)(i)Numbers.
(ii)Because it is a lot of numbers to choose from.
(12204)(i)I would take cm and always round up when it is X,5 and down when it is X,4.
(ii)I would say it is the most common way to do it so people all over the world would understand it.

Level 2These responses indicate a reasonable attempt at answering the question but with no explanation of how the answer was obtained or an explanation was given that was not related to the data or question. At this stage explanations resort to personal experience rather than referring to the data. For example:

(8115)(i)I would be particular, but close my eyes and put my pen down on a number.
(ii)It is a random selection and easy to do.
(8202)(i)26, 27, 28, 29
(ii)you don't see many people with that sized feet.
(8207)(i)29 cm
(ii)I don't know.

(9205)(i)26 to 30
(ii)Because these are the two numbers I felt like selecting.

Second Group

The second group of responses show an understanding of the question and attempt to rationalize the reduction of the data. These attempts to process the data are hampered by the lack of experiences and tools for reducing data. Again three levels were observed, coded as 3, 4 and 5. Sample responses are:

Level 3These responses indicate that, in attempting to justify a reasonable estimate or estimates, the reason did not refer to a feature of the data but to some feature of the question itself. For example:

(7213)(i)30 cm
(ii)If you rounded them up to the nearest 10 cm you would get 30 cm.
(10201)(i)28 and 30
(ii)Because that would tell you how big the foot was to the closest cm.

Level 4These responses indicate that the data was used to obtain a reasonable estimate or estimates and an awareness that the data needed to be used to justify this answer. However, restricted experiences at data reduction result in all data being quoted as necessary in the reason. For example:

(11111)(i)26, 27, 28, 29, 30, 31, 32, 33
(ii)every childs(sic) foot is the length of one of these numbers
(11201)(i)I would use the numbers 26 to 33.
(ii)This covers all the sizes of the left feet.

At this stage there is a divergence of the responses into TWO distinct paths which appear to develop at seemingly parallel rates. These are labelled

Path A for responses which reduce data based on measures of central tendency

Path B for responses which reduce data based on measures of dispersion.

Level 5 These responses indicate that the data could be reduced to a simple statistic. However, attempts to justify the answer usually result in an unsophisticated description of the statistic constructed. This indicates a readiness to engage in data reduction but a lack of experience and tools needed to produce a statistically sophisticated response. For example:

Path A

(8209)(i)I would select numbers 28, 29, 30
(ii)They are approximately the size of the students feet in the class.
(10202)(i)28 to 30
(ii)It is the average size for the students.
(11215)(i)29 cm
(ii)Because 29 is the average foot size.
(12208)(i)29
(ii)It is the average number (middle number)

Path B

(7203)(i)26 - 33
(ii)so you don't have to write so much
(11211)(i)The numbers, used to represent the size of the left feet in the class, that I would choose would be 26, 27, 28, 29, 30, 31, 32, 33.
(ii)The reason I chose these numbers is that if I put down one of every size measured it would give you an indication of the range (from shortest to longest).

Third Group

The final group of responses indicate an understanding of both the question and the data. The reduced form of the data is given in an acceptable statistical form and the explanation attempts to relate back to the data. The three levels of responses are coded as 6, 7 and 8, with the first two being split into A and B paths.

Level 6 These responses indicate that the data could be reduced to the form of a simple statistic and the reason for making such a selection related back to the data. However, the reason stated appears to reflect a taught definition rather than a true appreciation of the data.

Path A
Mode -
(10205)(i)28
(ii)Reason is because 28 happens the most.
(11107)(i)28 & 30
(ii)Because these 2 sizes are the 2 most common sizes in the results.
(11206)(i)30 cm because this is the mode left foot size.
(ii) It is the mode.
Median -
(10209)(i)29
(ii)It is the median
(11208)(i)30
(ii)In the middle.

Mean -
(10214)(i)28 and 29
(ii)The mean was 28.8
(10215)(i)28.86
(ii)because it is the mean of all the feet
(12113)(i)I would give all the numbers and divide by the number given to give an average (to nearest cm) size of the students feet.
(ii)I would do it this way because I believe it would be the fairest way to find the average size because the more numbers you take the closer and more accurate the survey will be.

Path B

(8210)(i)29 to 30 cm
(ii)because most feet are around that size.
(11103)(i)26 & 33
(ii)I gave the shortest & longest numbers to show the variation.
(12201)(i)26 - 33
(ii)That is the range.

Level 7These responses indicate the need to present more information by discussing more than one statistic. Reasons given still reflect the definitions of the statistics rather than discussions of properties of the data. Some responses discuss concepts which relate to statistical measures from the other path, for example a discussion of measures of central tendency may mention also some aspect of measures of dispersion. This is the stage where the two paths are beginning to converge again. The responses which use features of both paths are represented after some sample responses from Paths A and B.

Path A

(8106)(i)28 & 29
(ii)Because the no 28 is the most frequent one & also it is ruffly(sic) halfway between 26 & 33.
(9108)(i)I would chose the median or the mode of the feet
(ii)Because thats all I could think of
(9114)(i)30, 29
(ii)29, 30 is approximately the mean, median and mode of the numbers.
(11106)(i)The numbers I would select to represent the size of the left feet in that class would be 28 and 30
(ii)because they have the highest number in the class and those two are the average of them all.
(11213)(i)29 cm.
(ii)I would select this because it seems to be the average length of the people in the class. Also it is nearly the middle number.
(12207)(i)30 cm
(ii) because it worked out to be the average and also the most number of people in the class had their left foot 30 cm.

Path B

(9101)(i)I don't really understand this question but I'll say the top one the bottom one (length) and the middle one
(ii)it gives a fair idea of the range in sizes.

Responses which are not in Path A or B, data reduction involves the use of measures of both central tendency and dispersion.

(10106)(i)I would use 28 and 30.
(ii)Because most of the classes(sic) left feet were around that size and plus the sizes of 28 and 30 were most common.
(10213)(i)I would give the numbers 28 and 29
(ii)I would give these numbers as there are about the middle of the class and you can have a deviation of 3.
(11108)(i)26, 27, 28, 29, 30, 31, 32, 33 or an average
(ii)To show the range and the average.
(12210)(i)The mean or the average size of the left feet would represent the whole class.
(ii)The average gives an indication of the size foot of every student in the class. The standard deviation from the mean would also be useful in representing the class as this takes into consideration every student.

Level 8These responses indicate the use of both measures of central tendency and dispersion and the use of features of the data in an attempt to establish a link between the two.

(12209)(i)29
(ii)It is the middle number. The answers fluctuate a little to the negative and the positive sides of the number 29 (but mostly range around it).
(8215)(i)I would select either the mode 30 or the median 29
(ii)The mode occurs most often and would be the most common foot length in the class, 29 is the middle of the range in foot sizes and as such should come close to most of the non-extreme values.

The responses presented in each of the levels are only a sample of all responses but are considered to be representative of those given by students. It was possible to grade all responses into one of the nine levels presented and the results are now presented according to academic year, mathematical ability and gender.

The results arranged by academic year are presented in Table 1. These

data illustrate a number of interesting points. First, there are only two students (2%) from the three senior years whose responses fall within the first group (Levels 0,1 and 2), whereas in Years 7, 8 and 9

there are a number of students (17%) who have not fully understood the question. Second, of the four Level 8 responses, 3 were in Year 12 and 1 in Year 8. Third, there is a large bulge in all years at Level 6. Last, there are many more students whose responses reflect Part A than Path B, at least triple the number in every single academic year. This suggests a slight improvement in level of response with increasing academic year with a tendency to feel satisfied once a simple statistic has been used in the data reduction process. There is a definite preference for measures of central tendency rather than dispersion when describing the reduction of the data, irrespective of academic year.

Table 1
Response Level and Path by Academic Year

The results arranged by mathematical ability are presented in Table 2 for response level and Table 3 for response path. Testing the hypothesis that the level is independent of the mathematical ability of the student yielded c2 = 22.15 (6 d.f.) giving a probability of less than 1%. There is strong evidence to suggest that the coding level depends on the general mathematical ability of the student who gave the response. In the first group there are more low ability and less middle and high ability students than expected while in the Levels 7 to 8 there are far less low ability students than expected. Testing the hypothesis that the response path is independent of the mathematical ability of the student c2 = 3.00 (2 d.f.) giving a probability of more than 20%. This suggests that the coding path of any particular response is independent of the general mathematical ability of the student who gave the response.

Table 2
Response Level by Mathematical Ability

	Mathematical Ability			
Level	Low	Middle	High	Total
0-2	11	3	3	17
3-5	13	9	16	38
6	34	33	25	92
7-8	2	15	16	33
Total	60	60	60	180

Table 3
Response Path by Mathematical Ability

	Mathematical Ability			
Path	Low	Middle	High	Total

	A	34	4439	117
B	10	5	6	21
	Total	44	49	45 138

The results arranged by gender are presented in Table 4 for response level and Table 5 for response path. Testing the hypothesis that the level is independent of the gender of the student yielded x2 = 3.42 (3 d.f.) giving a probability of more than 30%. Testing for dependence

between the response path and gender of the student yielded x2 = 1.23 (1 d.f.) giving a probability of more than 10%. Hence it is reasonable to assume that both the coding level and path of any particular response is independent of the gender of the student who gave the response.

Table 4
Response Level by Gender

		Gender		
	Level	Female	Male	Total
	0-2	61117		
	3-5	172138		
6		474592		
	7-8	20	1333	
	Total	9090	180	

Table 5
Response Path by Gender

		Gender		
	Path	Female	Male	Total
	A	63 54	117	
B	8	13	21	
	Total	71 67	138	

These results suggest that, for the process of data reduction given raw data, the level of response improves progressively with academic year and many more students use measures of central tendency rather than dispersion to describe the data. The gender of a student does not appear to affect the response given but students with a higher mathematical ability are more likely to give a higher level response.

Analysis of Responses to Part II

The second part, II, was a general data reduction question in which the data was in a graphical form. The question was designed so that no

reference was made specifically to any measure of central tendency or dispersion. However, answering this question meant that students also needed to be able to understand and interpret the graph before they were able to engage in data reduction. The question, as it was presented to the students, is given in Figure 2.

Question Part II
A teacher was interested in how students performed in a spelling test to decide whether they needed extra help. The graph below represents the scores out of 10 achieved by the 28 year 9 students in the class in the spelling test.

(i) If you were asked to give a number, or numbers, which could be best used to represent the score in the spelling test of students in that class, what numbers, or numbers, would you select ?
(ii) Give reasons for your selection.

Figure 2

When the responses to Part II were analyzed they fell into similar levels to the Part I responses, except that no Level 3 responses were identified. The hierarchy of responses includes the same nine levels, arranged into the same three groups, as the analysis of Part I

responses.

It was possible to grade all of the responses into one of the nine Levels 0 through to 8 and the results were then considered according to academic year, mathematical ability and gender. The results arranged by academic year are presented in Table 6. From this data a number of interesting points can be observed. First there are less students (12%) from the three senior years who responses fall within the first group, than in Years 7, 8 and 9 where there are a number of students (20%) who have not even understood the question. Second, there was only one senior and no junior students whose responses were coded as Level 8. Third, there were no students in Level 3 and only 1 in Level 4. Fourth, there is a large bulge at Level 6 in all years. Last, there are approximately three times as many responses in Path A as in Path B. Although there are less seniors in the lower groups there is not much indication of an increase of level with academic year and irrespective of academic year many more students favour Path A than Path B.

Table 6
Response Level and Path by Academic Year

The testing of the hypothesis that the level is independent of the mathematical ability yielded $c2 = 21.90$ (6 d.f.) giving a probability of less than 1%. There is strong evidence to suggest that the coding

level depends on the mathematical ability of the student who gave the response. In the first group there are more low ability and less middle and high ability students than expected while in the Levels 7 to 8 there are far less low ability students than expected. Testing the hypothesis that the response path is independent of the mathematical ability of the student c2 = 1.45 (2 d.f.) giving a probability of more than 30%. This suggests that the coding path of any particular response is independent of the mathematical ability of the student who gave the response. The testing of the hypothesis that the level is independent of the gender of the student yielded c2 = 10.37 (3 d.f.) giving a probability of less than 2%. There is evidence to suggest that the level of the response depends on the gender of the student with more boys giving lower level responses and more girls in the higher levels. Testing for dependence between the response path and gender yielded c2 = 0.01 (1 d.f.) giving a probability of more than 90%. Hence it is reasonable to assume that the path of any particular response is independent of the gender of the student who gave the response.

For the process of data reduction, where data are given in a graphical form, the level of response improves progressively with academic year and many more students use measures of central tendency rather than dispersion to describe the data. This preference does not appear to be affected by either the gender or the mathematical ability of the student. Girls and students with a higher mathematical ability tend to give a higher level response when the data is presented as a graph. In the next section the two Parts, I and II, of the question are compared.

Comparison of Part I and Part II

The framework appears to be adequate in explaining students' understanding as far as the basic concepts of data reduction are concerned as it was possible to allocate all responses to a level. The slight upward shift in the trend of the responses over the academic years suggests the suitability of the levels as increased understanding would be anticipated over the years of secondary education as mentioned previously.

Some students appear to have found these questions difficult with many students being able to provide either no answer or a personal response. However, this group includes those who could not interpret or misinterpreted the graph. Comparing the numbers of students within each level and within each path for the two parts of the question assists with highlighting trends and differences.

Testing the hypothesis that the level is independent of the part of the question yielded c2 = 16.62 (6 d.f.) giving a probability of less than 2% and hence a result which is significant. This indicates that the level into which a response is coded is strongly dependent on whether Part I or Part II of the question is being coded. Many more responses

than expected were coded at the lower levels in Part II while Part I had more responses than expected in the uppermost levels. The students exhibited a higher level of understanding when the information was presented in a table rather than as a graph. When the number of responses that were graded into each path for Parts I and II were tested, the hypothesis that the path is independent of the question part yielded c2 = 2.71 (1 d.f.) giving a probability of approximately 10% and hence a result which is not significant. This indicates that the path into which a response is coded is independent of whether Part I or Part II of the question is being coded.

As far as the analysis of academic year is concerned, comparing the results presented in Tables 1 and 6 there are four noticeable trends. First, there are more junior students in the first group (Levels 0, 1 and 2) than there are senior students. Second, there are mainly senior students in Levels 7 and 8. Third, there is a large bulge in the numbers at Level 6 in every year. Fourth, there are many more Path A responses than Path B in every year. The general trend is for a slight increase in the level of performance of the students over the six academic years and a preference for measures of central tendency in the data reduction process.

The greater number of students who have responded by describing measures of central tendency (Path A) could perhaps be due to the heavy emphasis that so many teachers place on mean, mode and median. This gives the students a restricted set of experiences on which to base their data reduction and in fact may force students who would naturally be inclined to follow Path B processing to follow what to them may be a less natural reasoning pattern. The large bulge of responses at Level 6 is the point at which students are using one simple statistic to describe the data. This may be due to the fact that once students have been presented with simple statistical facts at school there are limited opportunities for data exploration and so they do not have the chance to develop more advanced data reduction skills.

Despite the similarities, there are some differences between the sets of data for the two parts of the question. Three in particular show up in a comparison of the results presented in Tables 1 and 6. First, the number of senior students on Level 0, 1 or 2 is much larger for Part II (11) than for Part I (2). Second, the overall number of responses at Level 7 or 8 is much larger for Part I, (33), than for Part II (18). Third, Levels 3 and 4 have far fewer students for Part II (1) than for Part I (11).

The small number of students in Levels 3 and 4 in Part II could be due to the fact that there were more responses in the lower group for that part. There were a number of students who completely misinterpreted the graph and this made it impossible for them to formulate a sensible answer for the question. These responses were graded in the first group and so would help to account for the larger number of students in the

lower levels for Part II. This is a problem associated with understanding the graph, rather than the reduction of data and appeared to particularly be a problem with Year 10 students. The larger number of higher level responses for Part I suggests that students are better able to make detailed statistical descriptions of data when they are presented as raw data rather than as a graph.

The performance of students appears to be affected in a similar fashion by varying mathematical ability for both parts of the question. There is strong evidence that the higher ability students are better able to understand the process of data reduction but the strong preference for the use of measures of central tendency rather than dispersion appears to apply across all mathematical abilities. This is the case whether the data are presented in raw form or as a graph.

The path chosen is not influenced by the gender of the student in either part of the question. There is a universal preference for a description based on measures of central tendency. However, gender influences the level of the response but only for Part II of the question. When the data are presented as a graph the girls are able to give a data reduction which is statistically superior to that given by boys. However, when presented with raw data the performance of boys and girls is similar.

From the above it appears that it is possible to develop a hierarchy of levels to categorize responses and the responses appear to fall into categories in an explainable fashion. Also, it appears that there are two possible paths of reasoning in the second and third groups and the form in which the data is presented may influence a student's choice of method of data reduction. Next is a comparison of the performances of individual students over the two parts of the question.

Comparison of Student Performances

The performance of students over the two parts of the question is now considered to determine whether there is consistency in the way in which any particular student responds when the data is presented in different forms. First the levels at which the student responded are compared. Then the paths along which the student modelled their response are compared.

To analyse the level of each student's response to Part I and Part II the information was condensed so that each student's grouping is considered rather than level to allow for analysis. When testing the hypothesis that the group into which a student's response fell is independent of the question part, $c2 = 54.38$ (4 d.f.) giving a probability of less than 0.1%. This result is highly significant strongly supporting the conclusion that there is a dependence of the

ranking of a student's response for one part of the question to the ranking of the response for the other part of the question. This is expected as 70% of students fall into the same group for the two parts of the question.

To investigate the relationship between the levels for each of the students on the two parts of the question a Nonparametric Sign Test was used. Of the 180 differences 89 were 0 and of the remaining 91 differences there were 57 which were positive. Testing the hypothesis that the codings are the same for the two parts of the question (that is, the proportion of positives is 0.5), Z = 2.31 giving a probability of approximately 2%. This means that the evidence against the hypothesis is significant and the proportion is not 0.5. As there were more positives than negatives it could be assumed that the students

tended to attain a higher coding on Part I of the question than on Part II. A student is likely to show more understanding of the data reduction processs when the data is presented in a raw form rather than as a graph.

Students were then compared to determine whether there is a link between the Path, A or B, taken for the two parts of the question. The categories Before, A , B and After are used to test the hypothesis, that the paths followed for Parts I and II of the question are independent, yielded c2 = 71.79 (9 d.f.) giving a probability of less than 0.1%. This evidence strongly supports a relationship between the path chosen in each of the two parts of the question. There is consistency in the processing path over the two parts of the question. In fact, if only the 117 students who were in one of the response paths for both parts of the question are considered then 101 (86%) of those students followed the same path in both parts (that is, both A or both B) while only 16 (14%) followed different paths.

These results show that there is not only a connection between the level or group of a student’s response on Parts I and II but also a connection between the path chosen when formulating the response. The the level for the classification of student responses over the two parts are dependent with data reduction being performed as a lower level when the data is presented as a graph than when it is in a raw form. The ability to read and interpret the graph is expected to have some bearing on this result. However, in attaining the level of response that they do the processing path taken by those students does not vary with the form of data presentation. Those students who have a natural inclination to consider measures of dispersion when reducing data may be more inclined to do so whether the data is presented in raw form or as a graph. Whereas those more inclined to use measures of central tendancy will do so irresepective of the form of presentation of the data.

The comparison of the grading of responses to Parts I and II indicate that the levels used to categorize the responses are sufficient to show trends which are not totally unexpected. These levels are now used to describe a framework which could be used to categorize students' responses to data reduction questions.

SOLO Taxonomy Framework

The levels described earlier are now used along with the SOLO Taxonomy to create a framework which could be used in future to code student responses. The first group of three levels contained responses which were ikonic in character while the second and third groups represented two different cycles in the concrete symbolic mode. Examples of responses in the following discussion are referred to by the subject's research number.

The responses in the ikonic mode suggest no link could be made between the required task and any sort of symbolic representation. Such responses were mainly from students in the junior years. Within this mode, there is a framework of growth with prestructural responses appearing as Level 0, when no attempt is made to answer.

The responses in Level 1 are a mixture of unistructural and multistructural responses. It was difficult to separate these responses because of the small overall number of responses in the ikonic mode and, because the students are not on task, it is difficult to determine how many visual cues of the question are in focus. These responses range from those like (7214) who answers the question literally

choosing "numbers" when asked to choose a number or numbers, through those who focus on irrelevant aspects of the question like (12204) who discusses rounding off the measurements, to those who discuss foot size but don't choose relevant data like (7102) who wants to use the odd numbers to represent the left feet because most people are right handed.

Finally, the relational level within this mode, corresponds to the Level 2 responses. Here the responses indicate that the various aspects of the question have been considered and understood sufficiently to produce a reasonable answer. However, attempts to justify the answer were usually unsuccessful or resulted in a personal comment. These range from (8207) who did not know why 29 cm would represent the foot size through (9205) who chose 26 to 30 because these were numbers he felt like selecting.

The responses in the second and third groups have been able to link the concepts in the question to concrete experience. Every indication is that the question has been understood and the reasons given link directly to the question or the data. These responses are in the

concrete symbolic mode. Within this mode two cycles became evident each containing three levels, unistructural (U), multistructural (M) and relational (R).

The first cycle involves appreciating that it is possible to reduce data into a more useable form. The elements in the first cycle are the actual pieces of data themselves (data items). A relational response in the first cycle is not achieved until all data items have been considered as a functioning set and represented in a concise form. In this cycle, the unistructural, multistructural and relational levels correspond to the levels 3, 4 and 5 as outlined earlier.

U1 - These responses present a simple answer as a reduction of the data but focus remains on the answer with the only reason given relating back to the question rather than to the data. A typical response is that of (7213) who used 28 and 30 cm to respresent the data but could only give as the reason “that would tell you how big the foot was”.
M1 -These responses still indicate a simple answer but attempts to focus away from this back to the data fail when attempting to combine the information. This results in quoting that all data are needed to represent the data. A typical responses are those of (11201) who chose 26 to 33 to cover all the sizes of the left feet.
R1 -These responses indicate that the answer and reason have combined the various pieces of data in the reduction process. This is the stage where there appears to be a split into two separate Paths, A and B, for the preferred styles of processing. A student who processes along Path A tends to focus in on the central tendency of the data during reduction while a student who processes along Path B focuses on the overall spread of the data. Responses in Path A range from those like (8209) who selected 28, 29 & 30 as these are approximately the size of students feet to those like (11215) who chose 29 as it is the average foot size. Responses in Path B range from those like (7203) who chose 26-30 for the foot size so as not to have to write too much.

The second cycle involves appreciating that the reduction of the data creates statistic(s) which are being used to describe the features or behaviour of the data. The elements in the second cycle are the various features (or properties) of the data which statistical data reduction are trying to describe. A relational response in the second cycle is not achieved until the student is able to consider various data features and the fact that these are related when it comes to

considering the overall data description. In this cycle, the unistructural, multistructural and relational levels correspond to the Levels 6, 7 and 8 as outlined earlier. These first two levels still contain the separate A and B processing paths. However, at the M2 level there are also responses which show evidence of elements from both paths.

U2 -These responses use a simple statistic to reduce the data but the reason given is usually just restating the definiton of the statistic rather than relating it to the data itself. Path A responses mention one measure of central tendency either by a description like (10205) who chose 28 for the foot size as it happens the most or by its correct statistical term like (10209) who chose 29 because it is the median. Path B responses mention one measure of dispersion either, e.g., by its correct statistical term like (12201) who chose 26-33 as the foot size because it is the range.
M2 -These responses consider more than one aspect of the data and give a number of statistics in an attempt to give a more accurate impression in the reduction process. Responses in Path A mention a number of measures of central tendency. Path B responses mention a number of measures of dispersion or give a more detailed description of the dispersion. Responses at the M2 level which incorporate elements from both paths are typified by (11108) who chooses both range and average to describe the foot size.
R2 -These responses reduce the data by presenting information concerning both central tendency and dispersion. Some attempt is also made to use one style of statistic to qualify the other. It is at this stage that the processing is no longer on either Path A or B path but integrates the features of the data. A typical response is (12209) who chose 29 to represent the foot size as it is the middle but also discusses how the other numbers fluctuate either side of the 29.

Conclusion

Overall, the understanding of data reduction appears to be better when the data is presented in raw form rather than as a graph. Students are far more likely to reduce the data using measures of central tendency than dispersion irrespective of the form of data presentation. Students with higher mathematical ability show more understanding of data reduction and when the data is presented as a graph girls show more understanding than boys. The framework presented above could be useful in determining the stage a student has reached in understanding data reduction. The use of the SOLO Taxonomy allowed for a broader intepretation of the levels identified and provided a clear context from which development of statistical ideas can be viewed.

References

Australian Education Council (AEC) (1991). A National Statement on Mathematics for Australian Schools, Carlton, Victoria: Curriculum Corporation.
Biggs, J. & Collis, K. (1982). Evaluating the Quality of Learning: the SOLO Taxonomy. New York: Academic Press.
Biggs, J. & Collis, K. (1991). Multimodal learning and the quality of intelligent behaviour. In H. Rowe (Ed.), Intelligence, Reconceptualization and Measurement (pp. 57-76). New Jersey: Laurence Erlbaum Assoc.

Holmes, P. (1980). Teaching Statistics 11-16, Berks, UK: Schools Council, Foulsham Education.
Pegg, J. (1992). Assessing students' understanding at the primary and secondary level in the mathematical sciences. In J. Izard & M. Stephens (Eds), Reshaping Assessment Practice: Assessment in the Mathematical Sciences Under Challenge (pp. 368-385). Melbourne:

Australian Council of Educational Research.