REACTIONS TO DATA: STUDENTS' UNDERSTANDING OF DATA INTERPRETATION

Chris Reading, University of New England, Australia

This paper reports on a study of high school students' understanding of statistics. Grade 7 to 12 students (n=180) completed an open-ended written response task about data description and prediction. Follow-up interviews were undertaken with 15 students. The SOLO Taxonomy and the Rasch Model were used to analyse the polychotomously categorised data. This analysis indicated student responses fell into two cycles within the concrete-symbolic mode, with two parallel paths of cognitive growth within these cycles. These findings suggest a need to modify approaches to teaching so as to better take into account the level of students' understanding of statistics.

INTRODUCTION

"Statistical thinking is a way of recognising that observations of the world are not always totally correct" (Rowntree, 1981, p.18). Just as there are varying degrees of 'correctness' in observations (data), there are also varying degrees or levels in the ability to understand data. The question then arises as to how to describe this developing understanding. Early attempts at considering students' understanding of statistics included the careful sequencing of tasks and hierarchies of levels of understanding within content structure (Reading, 1996, 51-63). Now, there is a growing trend, in educational research, to investigate understanding from the viewpoint of analysing student responses. Such investigations have been simplified by the development of the SOLO (Structure of the Observed Learning Outcome) Taxonomy, a tool used in a variety of fields (e.g. fractions, Watson, Campbell and Collis, 1996; science concepts, Panizzon and Pegg, 1997; geometry, Pegg, 1997).

SOLO TAXONOMY

The SOLO Taxonomy is a theoretical framework which attempts to provide a language for categorising the levels of students' responses at the various stages in the development of understanding (Biggs and Collis, 1980). The evolving structure of the SOLO Taxonomy, consisting of developmental modes and levels within each mode, was consolidated by Biggs and Collis (1991, 62-65). Relevant modes for the present discussion are: *ikonic* where actions are made more abstract by internal representation in some form with thought drawing heavily on imagery and being frequently affect laden and *concrete-symbolic* where there is a significant shift in abstraction from direct symbolisation of the world through oral language to symbol systems that apply to the experienced world. Levels which have been identified within each of these modes are:

Unistructural - focus on the relevant domain, and pick up one aspect to deal with
Multistructural - pick up more and more relevant features but show no integration
Relational - integrate the parts, the whole having a coherent structure and meaning.
More recent research has identified not just one but a number of cycles of unistructural-multistructural-relational (U-M-R) levels within the concrete-symbolic mode (Lidster, Watson, Collis and Pereira-Mendoza, 1996; Moritz, Watson and Collis, 1996).

RESEARCH DESIGN

Four broad areas of statistics were investigated in the study of students' understanding of statistics undertaken at an Australian rural secondary school. This discussion only deals with the data interpretation area. For details of the four areas used, and discussion of investigation results in the other three areas, see Reading (1996). Both male and female students (n=180) in Grades 7 to 12 were randomly selected from high, middle and low mathematical ability groups and given open-format questions with no time restrictions for completion. Figure 1 shows the relevant data interpretation question. Another question, with similar format but the data presented in graph form, was also given to the students to determine whether the form of presentation of the data would influence the students' ability to interpret the data.

A well known intersection in Armidale has had a number of serious accidents. The number of serious accidents was recorded for the last ten years.

1981	1982	1983	1984	1985	1986	1987	1988	1989	1990
2	4	5	9	3	6	7	10	4	6

(i) Describe any pattern that you can see in the data.
(ii) Approximately how many accidents would you expect in 1991? Why?
(iii) Suggest some years in the future (after 1990) when you think the number of accidents might exceed 8. Why did you select those years?

Figure 1. Data Interpretation Question

The responses were analysed both qualitatively and quantitatively. First, responses were grouped according to the degree of statistical sophistication exhibited overall in the three parts of the response. These groupings were further refined to develop a hierarchy of levels of understanding using the SOLO Taxonomy framework. Second, the influence of academic year, mathematical ability and gender on these levels was investigated. Chi-squared Contingency Table Tests were used to identify any effects of mathematical ability and gender but for the relationship between academic year and response ranking general trends were discussed, due to low cell numbers. Next, QUEST software (Adams and Khoo, 1993) was used to apply the Rasch model to estimate overall statistical understanding for each student based on their rankings in all questions over the four areas. From this analysis, tau (measure of the difficulty in moving from one level to the next) and threshold (estimate of understanding needed by a student to have a 50% chance of having a response coded at that level) values were used to provide quantitative evidence of the cycles identified within the modes. Finally, a longitudinal aspect was investigated by retesting the same students twelve months later, and structured interviews, with a small sample of students (15), were used to investigate the effect of probing and prompting on response level.

HIERARCHY OF LEVELS

Initially, the responses were ranked according to statistical sophistication into nine levels arranged into three broad groupings. The first group of responses dealt with only the requirements of the question with no consideration for, or processing of, the data. The second group indicated that, having understood the question, an attempt was made to understand the data with descriptions being non-statistical in nature. The third, and final, group of responses showed a readiness to describe the pattern in the data in a more acceptable statistical form, with predictions relating to the data. Using the SOLO Taxonomy to describe this hierarchy the responses fell into two modes of observed learning outcomes, the ikonic and the concrete-symbolic.

The ikonic mode responses, the first identified group, suggest that the students were unable to link the required task with any sort of symbolic representation. Within this mode, unistructural and multistructural responses, range from those students who focus on personal observations triggered by aspects of the question to those students who attempt predictions which appear to have no relation to the data. Relational level responses indicate that various aspects of the question have been considered when producing an answer, however, attempts to justify the predictions are either nonexistent or rely on personal comments rather than relating to the data.

Responses in the second and third groups form two U-M-R cycles within the concrete-symbolic mode. The first of these cycles (U1-M1-R1) involves describing the pattern, but usually in a form which makes prediction a difficult task. The elements for consideration are the actual pieces of data themselves (data items).

U1 - responses present either a simple pattern as an answer or infer a pattern from the explanation of the prediction, using one data item. Focus remains on the answer given and does not relate back to the question or the data, making it very difficult to make predictions.

M1 - responses still indicate a simple pattern as an answer, and, when focusing back to the data, attempts to combine more information into the description fail. This results in only an overall impression of the pattern being given. Answers suggest that more than one data item have been considered, but predictions still have personal comments as justification, rather than using the data or pattern. In attempting to describe the pattern many students at this level just quoted all data.

R1 - responses present a simple description of the pattern with discernible features which encompass all of the data. The predictions show evidence of the pattern having been used to produce them, but an explanation of how the pattern was used could not be made. An interesting feature of these responses was the split into two parallel paths. One, referred to as Path A, using features based on central tendency and the other, referred to as B, using the overall spread (dispersion) of the data to describe features of the pattern.

Responses from the second cycle (U2-M2-R2) involve appreciating that the pattern in the data can be used to justify predictions. The elements to be combined in the second cycle are the various features of the data which constitute the pattern.

U2 - responses give a simple description of the pattern in the data and use this pattern to make predictions. At this level, difficulty is encountered in justifying the prediction using the pattern, and the explanation may just quote that the prediction follows the pattern. Responses were identified in both Path A and B.

M2 - responses indicate that more than one feature of the data have been considered when describing the pattern, and attempts to justify the predictions use features of the pattern. Responses were identified again in both Path A and B but there were also responses which incorporate elements from both paths.

R2 -responses present features of the pattern concerning both central tendency and dispersion and accurately justify the predictions using the pattern described. It is at this stage that there is integration of the features of the data.

FACTORS INFLUENCING RANKINGS

The level of the response appeared to increase gradually with increasing academic year. Only 2 students (3%) from Years 11 and 12 (seniors) while 19 (32%) Year 7 and 8 students (juniors) had responses ranked in the first group. On the other hand, 5% of juniors compared to 40% of seniors responded at the M2-R2 levels. Overall, more than half the student responses were at the R1 and U2 levels, with no Year 7 response above U2 and no Year 12 response below M1. About twice as many student responses fell into Path B than into Path A, in all Years, except 10 where the proportion was equal. These results suggest that there is some improvement in understanding from Year 7 to Year 12 and many more students prefer to make inferences based on dispersion concepts.

Both mathematical ability and gender showed significant association with the level of response ranking. There was strong evidence ($\chi^2 = 26.17$, 6 d.f., $p < 0.001$) to suggest that coding level depends on mathematical ability. In the first six levels there are more low mathematical ability and less middle and high mathematical ability students than expected, with the reverse in the third group. A significant result ($\chi^2 = 13.12$, 5 d.f., $p < 0.05$) also suggests that the level of the response depends on gender. There are more responses given by females than expected in the third group. However, there is evidence to suggest that coding path is independent of both mathematical ability ($\chi^2 = 0.48$, 2 d.f., $p > 0.70$) and gender ($\chi^2 = 1.15$, 1 d.f., $p > 0.20$).

The form of presentation of the data, as a graph rather than in raw form, does not appear to have had a marked effect on assessing students' level of understanding. Similar hierarchy levels were found to be appropriate for ranking with the quality of response increasing with academic year. Results, as far as mathematical ability and gender are concerned, were consistent with those for the raw data question except that the level of response did not depend on gender. When the data were presented in graph form this appeared to help the males to achieve a higher level response than with the raw data and hence perform at a similar level to the females.

RASCH ANALYSIS

The Rasch analysis reinforced the notion that the coding of the responses fell into two distinct cycles in the concrete-symbolic mode. Tau values suggested that students found it difficult to move into the first cycle of the concrete-symbolic mode, using the data to construct their answer. The easiest steps appear to be moving into the latter part of this first cycle (M1 and R1), that is, once students have been able to recognise a simple pattern in the data, they find it relatively easy to describe that pattern or make predictions in simple terms. Difficulty increased again moving into the second cycle, where the data were used to justify predictions. Students experienced greater difficulty attaining the upper levels in the second cycle (M2 and R2) which does not appear to be consistent with the first cycle pattern but can possibly be explained by the fact that there were few responses at these levels and secondary students found it quite difficult to construct this level response.

Thresholds values clearly indicated the existence of a first cycle with greater increases in the level of understanding required to enter and exit the cycle and easier access for students within the cycle. However, again, the structure of the second cycle is not so clear with the steps within the cycle showing a great increase in level of understanding required. Further research, probably requiring assessment of tertiary students, may help to clarify the existence and character of this second cycle in the concrete-symbolic mode.

LONGITUDINAL ASPECTS

In the longitudinal retesting, the previously identified levels of understanding proved sufficient to code the responses and subsequent analysis substantiated the initial results. However, gender influence on response level was found to be not significant on retesting, and, as previous significance was not strong, it is possible that gender influence is not so important as mathematical ability which still had a strong significance in the retesting.

The correlation of 1991 and 1992 student estimates of understanding was 0.61 (115 d.f.) which is very significant ($p < 0.001$) indicating a consistency in performance (positive correlation showing that higher estimates are linked for each year). Change in overall statistical understanding (mean = -0.03) was not significant ($t = -0.76$, 116 d.f., $p > 0.05$) over the twelve months, indicating that over the twelve-month period there does not appear to be an increase or decrease in statistical understanding.

PROMPTING AND PROBING

Most students added more to their initial response when probed and prompted in an interview situation but in many instances this gave a more detailed answer still at the same level rather than increasing the level of the response. Probing allowed one third of students to increase their level, but, when pressured, students sometimes reverted to lower levels for support, thus exhibiting a reliance on the ikonic mode in justifying the initial concrete-symbolic response. Prompting produced an increase in response level for half of the students. An unexpected and noticeable effect of prompting was a path change in the new response. Changes from Path A to Path B responses, or B to A, suggests that students have the facility available to respond along either path but will use their preferred path unless pressured to delve further.

CONCLUSIONS

These findings suggest a need to modify approaches to teaching so as to better take into account the level of students' understanding of statistics. The levels and modes identified could be useful in determining the stage a student has reached in understanding the process of data interpretation. Ikonic mode responses tend to rely on personal experience, rather than data, to make predictions. Once a response exhibits the realisation that predictions need to be substantiated, it indicates that the transition to the concrete-symbolic mode has been made. The results indicate that most students make this transition from ikonic to concrete symbolic, learning to make use of data, during their early secondary schooling but not all become comfortable with using data to justify their discussions by the end of secondary schooling. Caution should be used when teaching students with low mathematical ability as they appear to have greater difficulty with using data to justify their discussions. The identification of two distinct paths suggests that teachers should vary their data discussion presentations to include dispersion as well as central tendency so as to cater for all students.

REFERENCES

Adams, R. J. and Khoo, S. T. (1993). *QUEST The Interactive Test Analysis System..* Australian Council for Education Research Limited, Hawthorn, Victoria.

Biggs, J. B. and Collis, K. F. (1980). The SOLO Taxonomy. *Education News*, *17(5)*, 19-23.

Biggs, J. B. and Collis, K. F. (1991_. Multimodal Learning and the Quality of Intelligent Behaviour. In Rowe, H. A. H. (Ed.), *Intelligence: Reconceptualisation and Measurement,* Hillsdale. N.J., Lawrence Erlbaum, 27-76.

Lidster, S.T., Watson, J. M., Collis, K. F. and Pereira-Mendoza, L. (1996). The Relationship of the Concept of Fair in the Construction of Probabilistic Understanding. In Clarkson, P. C. (Ed.), *Technology in Mathematics Education*, *Proceedings of the 19th annual conference of the Mathematics Education Research Group of Australasia, July, 1996*, Melbourne, 352-359.

Moritz, J. B., Watson, J. M. and Collis, K. F. (1996). Odds: Chance Measurement in Three Contexts. In Clarkson, P. C. (Ed.), *Technology in Mathematics Education, Proceedings of the 19th annual conference of the Mathematics Education Research Group of Australasia, July, 1996*, Melbourne, 390-397.

Panizzon, D. and Pegg, J. (1997). Investigating Students' Understandings of Diffusion and Osmosis: A Post Piagetian Analysis. Paper presented at the 1997 annual conference of the Australian Association for Research in Education, December, 1997, Brisbane.

Pegg, J. (1997). Broadening the Descriptors of van Hiele's Levels 2 and 3. in Biddulph, F. and Carr, K. (Eds) *People in Mathematics Education*, *Proceedings of the 20th annual conference of the Mathematics Education Research Group of Australasia, July, 1997,* Rotorua, 391-396

Reading, C. (1996). *An investigation into Students' Understanding of Statistics*., Unpublished Doctoral Dissertation, University of New England, Armidale, Australia.

Rowntree, D. (1981). *Statistics Without Tears*. Penguin Books.

Watson, J. M., Campbell, K. J. and Collis, K. F. (1996). Fairness and Fractions in Early Childhood. In Clarkson, P. C. (Ed.), *Technology in Mathematics Education*, *Proceedings of the 19th annual conference of the Mathematics Education Research Group of Australasia, July, 1996,* Melbourne, 588-595