INITIAL CONSIDERATIONS CONCERNING THE UNDERSTANDING OF PROBABILISTIC AND STATISTICAL CONCEPTS IN AUSTRALIAN STUDENTS

JANE M. WATSON AND KEVIN F. COLLIS
University of Tasmania

Following the questions raised by Watson (1992) concerning research in probability and statistics education in Australia in the 1990s, this paper reports on the initial trialling of items with 64 Grade 9 students. The analysis supports the belief that misconceptions observed in other countries also are present in Australia. Further, the application of a developmental cognitive model offers promise for classifying responses to items and structuring remediation procedures.

The rationale for a growing interest in research in probability and statistics education is given in Watson (1992). Briefly, topics associated with chance and data constitute 20 percent of the content of A National Statement on Mathematics for Australian Schools (Australian Education Council, 1991). As the newest areas in the mathematics curriculum, chance and data are least likely to be well understood by teachers at all levels. They are also the areas about which the least is known of students' understandings and misconceptions. As state curricula are amended to take account of the National Statement and professional development is planned to assist teachers, it is important to provide as much support as possible based on analyses of students' and teachers' understanding of these topics.

Watson (1992) also briefly reviews the history of overseas research into probabilistic and statistical understandings and offers some starting points for Australian research. These include studies of student beliefs about probability (e.g., Fischbein, Nello, & Marino, 1991; Garfield & delMas, 1991), conceptions of randomness (e.g., Konold, Lohmeier, Pollatsek, Well, Falk, & Lipson, 1991), sampling in relation to sample size and representativeness (e.g., Gal, Rothchild, & Wagner, 1989), the relationship between two variables (e.g., Konold, 1991), the applicability of measures of central tendency (e.g., Gal, Rothchild, & Wagner, 1990), and the understanding of conditional probability (e.g., Pollatsek, Well, Konold, & Hardman, 1987).

A theoretical basis for the analysis of student understanding could come from various sources. A developmental model, such as that suggested and extended by Biggs and Collis (1982, 1989, 1991) has proved useful in other areas of mathematics (e.g., volume measurement (Campbell, Watson, & Collis, in press) and fractions (Watson, Campbell, & Collis, in press)) and parallels the model suggested by Shaughnessy (1992). Two aspects of the Biggs and Collis SOLO (Structure of Observed Learning Outcome) Taxonomy will be exemplified in the analyses presented in this discussion. First is the unistructural-multistructural- relational (UMR) cyclic structure which operates within the modes of functioning. Of particular interest here are the concrete symbolic mode where most school learning occurs and the ikonic mode which commences functioning earlier and is associated with feelings, intuitions and imaging. Second is the multimodal aspect of functioning which recognises various ways in which interactions take place between modes, in particular the ikonic and concrete symbolic, as both continue to develop throughout childhood and adolescence.

The remainder of the discussion will focus on the initial sampling of student responses to explore conceptions suspected to be of interest and on the usefulness of the SOLO model in analysing results.

ITEMS

Items were selected and adapted from those reported by Fischbein and Gazit (1984) on intuitive probability (six items), Fischbein et al., (1991) on outcomes and sample spaces (three items), Reading and Pegg (1992) on sample selection (two items), and Garfield and delMas (1991) on conceptions of probability (three items). Following the interest of Mokros and Russell (1992), an item asked for an explanation of the term 'average'. Finally a problem solving situation was devised to explore student understanding of criteria which could be used to compare two data sets. Although administered differently the final section was based on ideas suggested by Wagner and Gal (1991) and Garfield (1992). The 12 items to be summarized in this paper are presented in full in Watson (1993); three are shown in Figure 1.

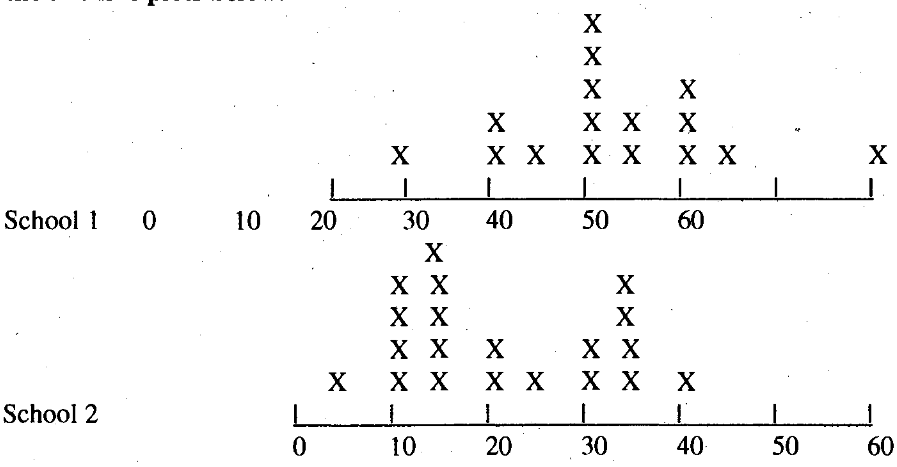
THE SAMPLE

Because the purposes of this preliminary study were to assess the items chosen, to explore possible misconceptions, to assess the potential of the theoretical developmental model, and to suggest strategies for a larger scale study, convenience samples were used. They consisted of two Grade 9 mathematics classes from Hobart, Tasmania, suburbs, totalling 64 students. They were 13 or 14 years of age and the sample was evenly split between females and males; gender differences are not explored here. Students completed the written questionnaire during class time. One of the classes had earlier in the year studied a unit of probability while the other had studied one on statistics. Although these would not have lasted more than a few weeks, it was expected that the sample would present some good responses to the items.

MISCONCEPTIONS

The 12 items used for this analysis led to an examination of the seven misconceptions listed in Table 1. The number of items associated with each are noted and the percentage of students exhibiting the misconception for the items is given. The percentages are estimates based on the number of responses to each item, ranging from 64 to 32.

Item A **Two maths classes, one from your school and one from the nearby high school, have had a competition to see which is better at quick recall of maths facts. The scores of the two classes are shown on the two line plots below.**



How would you analyse these data to show which school is better and which school would probably be your school?

Item B **Consider rolling two dice. Is it more likely to obtain**

(a) a 5 with one die and a 6 with the other?, or

(b) 6 with both dice?, or

(c) is the probability the same in both cases?

Item C **Again consider rolling two dice. Which is more probable?**

(a) To obtain the same number with both dice., or

(b) To obtain different numbers on the dice., or

(c) is the probability of each event the same?

Figure 1 Three items used in preliminary study

Hence non responses are ignored and this means the percentages are lower bound estimates for this sample. The percentages in Table 1 are also lower bounds in the sense that not all other answers gave correct, or well-expressed, responses - they just did not show the misconception indicated in the right hand column.

No students believed in lucky numbers and very few agreed with superstitions associated with putting a certain foot into a room first or with winning a chance game due to being older. Approximately a quarter of students displayed no understanding of average and a similar fraction used the total of scores to compare two groups despite differing sample sizes. Between 20% and 50% of students had difficulty using proportion to make judgements in relation to samples and populations. Nearly half had difficulty with the concept of randomness in relation to lottery numbers and/or sequences of binomial trials. Finally up to 85% of students had problems with the sample space when two dice are tossed.

Construct	Number of Items	Percentage difficulty for item
Luck (lucky numbers)	1	0%
Superstition (right foot, age)	2	5%, 2%
Average (no basic understanding)	1	24%
Compare different sized groups (using total)	1(A)	28%
Proportion (samples and populations)	3	26%, 20%, 56%
Randomness (lottery, births)	3	46%, 47%
Sample spaces (dice)	2(B,C)	83%, 35%

Table 1. Percentages displaying basic misconceptions in Grade 9.

ANALYSIS AND DISCUSSION WITH COGNITIVE MODEL

There are two ways that the SOLO framework can be used initially in a study such as this. First the items themselves can be classified by the level of response required to obtain a correct answer. Second responses can be categorised by the level actually exhibited by the student.

Consider first the constructs listed in Table 1. The first two, labelled Luck and Superstition, relate to intuitions which are associated with the ikonic mode of functioning. As the mode which begins functioning earlier, it would be expected that by Grade 9, students functioning well in the concrete symbolic mode would be expected to have rejected these intuitive level concepts as adequate responses. This would appear to have been the case for this sample of students. The other five constructs listed in Table 1 are related directly to operations of the concrete symbolic mode as they rely mainly on school-based instruction. They are listed in order of increasing difficulty as judged by the percentages showing the misconceptions. This seems to be a good starting point for suggesting a hierarchy of difficulty, as well as for indicating the functioning level required for a correct concept to be displayed. It is too early, however, to put much weight on such an ordering, particularly in the light of differing formats for the items.

When moving to the responses themselves, it is possible that ikonic and concrete symbolic functioning will be demonstrated on the same item, regardless of what is hypothesized to be required for a 'correct' response. In looking at the responses reported, from a cognitive perspective, two aspects are significant. One is the developmental structure of the response and the other is evidence of multimodal functioning. In the latter case, in this study, this means the interaction between intuitive/ikonic and concrete symbolic functioning.

For the purposes of this study the following SOLO notation is used to classify different levels of response for each item within the concrete symbolic mode:

U - uses one piece of relevant information to respond;

M - uses more than one piece of relevant information: these pieces of information to be of a similar type;

R - relates more than one piece of relevant information, of different types, in a coherent fashion.

In previous research on understanding of volume measurement (Campbell, et al., in press) and fractions (Watson, et al., in press), it was found that two UMR cycles operated in problems involving these constructs. The first cycle, labelled U_1-M_1-R_1, led to the correct understanding of the concept, while a second, labelled U_2-M_2-R_2, resulted when the concepts were applied in problem solving contexts of varying difficulty.

An application of this model is shown below in relation to the item concerning 'average'.

Response: *'Normal'*

This is classified as unistructural, U_1, as it shows a lack of consideration for the mathematical aspects of the construct, while selecting a descriptive common sense aspect.

Response: *'Average means the middle number'*

This appears multistructural, M_1, in that, in the context of the arithmetic mean, it implicitly considers a set of numbers and makes a judgement about a single number describing it. (Note the respondent was not aware of the median concept.)

Response: *'Average is when you add up your total digits and then divide it by the number of digits you have,'* and
'Each time you did something and it had answers that were a little bit apart you could find the average witch [sic.] would get it down to one number'

These would appear to exhibit a relational, R_1, understanding of the arithmetic mean as average.

Moving to item A in Figure 1, the question arises as to the application of this average concept to a given problem. Because a comparison must be made, it would be expected that a relational, R_2, response is required. The following response qualifies as an R_2 for this concept:

If I were to show which school was better I would find out the averages of both schools.
The higher the average determines which school is better.

The nature of the question precluded U_2 & M_2 responses in this item which indicates that we will need to redevelop the item to get the full range of responses. A U2 response would be expected to be similar to an R1 response in that the student would be able to calculate the 'average' (arithmetic mean) when asked to do so with data in a suitably obvious form. An M2 response would involve, say, the calculation of several means from given data and then a separate judgement being made for example on the largest or smallest. The R2 example given requires the student to make the decision that the arithmetic mean is the appropriate construct to use, calculate them for each group and then compare to reach a conclusion; in other words it requires an overview of the problem which is not required at U2 or M2 response levels.

Another example of functioning in the second UMR cycle is found in items B and C in Figure 1. As was found on an earlier item not discussed here, students had a good grasp of the sample space associated with a single die and could distinguish impossible, possible and certain events. This is representative of R_1 thinking in relation to the concept of sample space in this context. Items B and C, however, require an R_2 level of functioning in the concrete symbolic mode as they involve an understanding of the sample space in relation to throwing dice extended to the more complex situations where two dice are involved. The U2 level of response, represented by the following:

'You have 1 in 6 chances of getting a five and it is the same with a six, so they are the same'
and
'because there is a 50/50 chance of either'

These show an R_1 concept and justify the same probability for throwing a 5 and a 6, as two 6's, despite using two dice. An early R_2 response,

'Because you can get a 5 or 6 on either dice,'

justifies the higher probability for a 5 and a 6 by applying the concept over the extended sample space of the pairs of numbers involved.

For item C, where more correct answers were obtained, the justifications for 'the same' probability were still at the U_2 level:

'Because rolling dice is random. Any no could come up.'

'This is because the dice is fair and any sequence should be just as likely as any other.'

The following was judged M_2 because, although it has the right idea, it does not tie the concepts together in a coherent fashion:

'Because the dice has many different numbers but only one of each.'

The following response to the item seems to have related the concepts together satisfactorily to achieve the correct answer at the the R_2 level:

'Because there are more pairs of different numbers than there are pairs of the same number. So the chance of getting a double is smaller.'

As mentioned earlier, ikonic functioning may or may not take place in conjunction with concrete symbolic functioning. In the examples presented here the ikonic responses generally relate to beliefs and intuition, not to imaging. These responses are most commonly found in explanations where the students appear unable to reason effectively in the concrete symbolic mode. Hence for item B,

'Nobody knows what will come up on the dice - its [sic.] just fate'

and

'It occurs by chance only'

would be considered ikonic (IK) responses as would,

'It is all to do with luck',

in item C.

A response such as,

'I think it is a chance thing and you would have better odds getting random numbers',

on an item about a straight sequence of numbers in a lottery, shows lack of appreciation of chance in the context but has some idea of what odds means. It appears to be an IK-M_1 response, struggling with the transition from intuition to an understanding of the association of randomness and chance. Sometimes, however, there seems to be an interaction of intuitive (IK) and concrete symbolic reasoning at a higher level. The following discussion, for example, associated with the belief that entering a room with one's right foot first will enable a good performance, shows an appreciation of superstitions but the student makes the decision on the basis of concrete symbolic reasoning.

I don't agree with this [as it is not logical]. *But if Joseph believes that putting his right foot forward [helps] then that [is] what he should [do] because it is something which he believes in...'*

The response appears to show a high level of ikonic interaction with concrete symbolic reasoning and could be categorised as IK-U_2 in this context.

Although it has been possible to explore responses from those collected in the initial sampling of students which seem to exemplify the UMR cycle and multimodal functioning, it is clear that a great deal more work needs to go into item development to ensure that the items are flexible enough to allow for a full range of responses both within and between modes of functioning. It may be useful to devise items similar to those designed for the Collis-Romberg (1992) Problem Solving Profiles; such items would have the additional advantage that they can be used later to identify achievement levels to assist teachers in planning instruction and remediation.

ACKNOWLEDGMENT

The authors would like to acknowledge the assistance of Fiona Shalley in recording and collating the data in this study.

REFERENCES

Australian Education Council. (1991). A National Statement on Mathematics for Australian Schools. Canberra: Author.

Biggs, J.B. & Collis, K.F. (1982). Evaluating the quality of learning: The SOLO Taxonomy. New York: Academic Press.

Biggs, J.B. & Collis, K.F. (1989). Towards a model of school-based curriculum development and assessment using the SOLO Taxonomy. Australian Journal of Education, 33, 151-163.

Biggs, J.B. & Collis, K.F. (1991). Multimodal learning and the quality of intelligent behaviour. In H. Rowe (Ed.), Intelligence: Reconceptualisation and measurement. Hillsdale, N.J.: Lawrence Erlbaum.

Campbell, K.J., Watson, J.M. & Collis, K.F. (in press). Volume measurement and intellectual development. Journal of Structural Learning and Intelligent Systems.

Collis, K.F., & Romberg, T.A. (1992). Collis-Romberg Mathematical Problem solving Profiles. Hawthorne, Vic.: Australian Council for Educational Research.

Fischbein, E., & Gazit, A. (1984). Does the teaching of probability improve probabilistic intuitions? Educational Studies in Mathematics, 15, 1-24.

Fischbein, E., Nello, M.S., & Marino, M.S. (1991). Factors affecting probabilistic judgements in children and adolescents. Educational Studies in Mathematics, 22, 523-549.

Gal, I., Rothchild, K., & Wagner, D.A. (1989, April). Which group is better?: The development of statistical reasoning in elementary school children. Paper presented at the Society for Research in Child Development meeting, Kansas City, MO.

Gal, I., Rothchild, K., & Wagner, D.A. (1990, April). Statistical concepts and statistical reasoning in school children: Convergence or divergence. Paper presented at the American Educational Research Association conference, Boston.

Garfield, J.B. (1992). Personal communication.

Garfield, J.B., & delMas, R. (1991). Students' conceptions of probability. In D. Vere-Jones (Ed.), Proceedings of the third international conference on teaching statistics. Volume 1: School and general issues (pp. 340-349). Voorburg: International Statistical Institute.

Konold, C. (1991). Chance Plus: A computer-based curriculum for probability and statistics. (Annual Review: Year 2, NSF Grant No. MDR-8954626). Amherst, MA: Scientific Reasoning Research Institute.

Konold, C., Lohmeier, J., Pollatsek, A., Well, A., Falk, R., & Lipson, A. (1991). Novices views on randomness. In R.G. Underhill (Ed.), Proceedings of the thirteenth annual meeting, North American Chapter of the International Group for the Psychology of Mathematics Education, Volume 1 (pp. 167-173). Blacksburg, VA: VPI & SU.

Mokros, J.R., & Russell, S.J. (1992). Children's concepts of average and representativeness (Working Paper 4-92). Cambridge, MA: TERC.

Pollatsek, A., Well, A.D., Konold, C., & Hardiman, P. (1987). Understanding conditional probabilities. Organizational Behavior and Human Decision Processes, 40, 255-269.

Reading, C. & Pegg, J. (1992). Personal communication.

Shaughnessy, J.M. (1992). Research in probability and statistics: Reflections and directions. In D.A. Grouws (Ed.), Handbook of research on mathematics teaching and learning (pp. 465-494). New York: MacMillan.

Wagner, D.A., & Gal, I. (1991). Project STARC: Acquisition of statistical reasoning in children. (Annual Report: Year 1, NSF Grant No. MDR90-50006). Philadelphia, PA: Literacy Research Center, University of Pennsylvania.

Watson, J.M. (1992) What research is needed in probability and statistics education in Australia in the 1990s? In B. Southwell, B. Perry, & K. Owens (Eds.), Proceedings of the Fifteenth Annual Conference of the Mathematics Education Research Group of Australasia (pp. 556-567). Kingswood, NSW: MERGA.

Watson, J.M. (1993). Initial considerations concerning the understanding of probabilistic and statistical concepts in Australian students: Working Paper. Unpublished manuscript, University of Tasmania, School of Education, Hobart.

Watson, J.M., Campbell, K.J., & Collis, K.F. (in press). Multimodal functioning in understanding fractions. Journal of Mathematical Behavior.