Tracking Structural Development Through Data Modelling in Highly Able Grade 1 Students

Joanne Mulligan
Macquarie University
<joanne.mulligan@mq.edu.au>

Kerry Hodge
Macquarie University
<kerry.hodge@mq.edu.au>

Mike Mitchelmore
Macquarie University
<mike.mitchelmore@mq.edu.au>

Lyn English
Queensland University of Technology
<l.english@qut.edu.au>

A 3-year longitudinal study *Transforming Children's Mathematical and Scientific Development* integrates, through data modelling, a pedagogical approach focused on mathematical patterns and structural relationships with learning in science. As part of this study, a purposive sample of 21 highly able Grade 1 students was engaged in an innovative data modelling program. In the majority of students, representational development was observed. Their complex graphs depicting categorical and continuous data revealed a high level of structure and enabled identification of structural features critical to this development.

Our studies of early mathematical development have focused on the role of pattern and structure in promoting early algebraic thinking and generalisation (Papic, Mulligan & Mitchelmore, 2011; Mulligan, English, Mitchelmore, & Crevensten, in press). Several of these studies have shown that the provision of explicit and connected structured tasks over a school year can significantly advance the development of such mathematical processes as patterning and unitising, spatial structuring, multiplicative reasoning, and generalisation.

Aligned with these studies is the investigation of young children's data modelling and representational skills (English, 2012). With a growing emphasis on statistical reasoning in the curriculum, an integrated approach to the study of data modelling can bring much insight into young children's mathematical and scientific reasoning. This paper reports an exploratory study of how an emphasis on structuring graphical representation might enhance the development of metarepresentational competence (English, 2012). A central aim of the study was to identify explicitly the structural features that play an essential role in this development.

Background to the study

Over the past decade, the Pattern and Structure Project at Macquarie University has aimed to find reliable methods for measuring and describing the structural growth in mathematics of 4 to 8 year olds. In a series of studies, the project has promoted a strong foundation for mathematical development by focusing on critical, underlying features of mathematics learning. The construct Awareness of Mathematical Pattern and Structure (AMPS) was developed, which our studies have shown generalises across early mathematical concepts, can be reliably measured using the Pattern and Structure Assessment (PASA) interview, and is correlated with mathematical understanding (Mulligan & Mitchelmore, 2009). AMPS has two components: not only an understanding of common mathematical structures but also a tendency to look for patterns in new situations. Our belief is that a focus on the development of AMPS can bring increased coherence to the study of learning and teaching of early mathematics. The Pattern and

In V. Steinle, L. Ball & C. Bardini (Eds.), Mathematics education: Yesterday, today and tomorrow (Proceedings of the 36th annual conference of the Mathematics Education Research Group of Australasia). Melbourne, VIC: MERGA.

Structure Mathematical Awareness Program (PASMAP) was developed concurrently with the studies of AMPS to implement a corresponding pedagogy.

A large-scale two-year longitudinal evaluation study *Reconceptualising Early Mathematics Learning* found that Kindergarten students who engaged in the PASMAP program achieved significantly higher AMPS in the middle of Grade 1 than a comparison group (Mulligan et al., in press). A subsequent longitudinal project *Transforming Children's Mathematical and Scientific Development*[1] *2011-2013* extends the initial study with some students tracked through from 2010. This project explores the role of pattern and structure in mathematics and science learning in Grades 1 to 3 and tracks a new cohort of highly able students from Kindergarten through to Grade 2. This research integrates English's research on data modelling (English, 2012) with the study of structure.

Research on Data Modelling

Data modelling is a developmental process that begins with young children's inquiries and investigations of meaningful phenomena (e.g., exploring the growth of flowering bulbs under different conditions), progressing to deciding on the factors that may have influenced the observations; and then structuring, organising, analysing, visualising and representing data; organising and displaying data in simple tables, graphs and diagrams; and analysing the data to identify any relationships or trends (Lehrer & Schauble, 2005).

A longitudinal study of data modelling in Grade 1 (English, 2010) indicated that children as young as 6 years old can successfully collect, represent, interpret, communicate, and argue about the structure of data provided that they address familiar themes. The study has revealed young children's competence in structuring and representing data they collect, together with skills in creating multiple and varied representations, in spite of minimal instruction. An understanding that their different representations show the same data has also been evident in children's responses. Children's use of various inscriptions including labelling vertical and horizontal axes of their bar graphs and the sectors in their circle graphs, and their application of appropriate scaling and colouring, illustrate their awareness of the importance of constructing data representations that are readily interpretable. Overcoming obstacles and limitations in their representational material, such as using two columns to display extended data and adjusting their scaling to accommodate a wide range of values, indicates critical and flexible thinking in their dealing with data. Innovative representations, such as applying an awareness of the structure of a heart rate monitor graph, show young children's creative links with their world and their ability to reason analogically (English, 2012).

Theoretical Approach

Research into students' understanding of graphing has described four increasingly sophisticated stages of development (Prestructural, Unistructural, Multistructural, and Relational) ranging from simple showing attributes, methods of displaying data, to understanding relationships and variation (Watson & Fitzallen, 2010; Watson & Kelly, 2005). Prestructural representations lacked mathematical features such as countable items. At Unistructural level items were organised so that they could be counted, but at the Multistructural level arrangements were more structured, using grid-like features, for example. At the Relational level data revealed horizontal and vertical features that were

[1] Australian Research Council Discovery Project DP110103586 (2011-2013)

coordinated and ordered. To some extent, the development of our stages of structural development based on AMPS reflects the work of Watson and colleagues.

Our approach is to apply the AMPS construct to the analysis of students' development of data-structuring skills. AMPS is measured by assessing students' responses to a wide variety of mathematical tasks in terms of the critical features of the underlying structures. Responses are classified into the following five levels (Mulligan & Mitchelmore, 2009):

1. *Prestructural*. Children pick on particular features that appeal to them but are often irrelevant to the underlying mathematical concept.
2. *Emergent.* Children recognise some relevant features but are unable to organise them appropriately.
3. *Partial structural.* Children recognise most relevant features of the structure, but their representations are inaccurate or incomplete.
4. *Structural.* Children correctly represent the given structure.
5. *Advanced.* Children recognise the generality of the structure.

Applying these structural levels to the analysis of students' graphical representations requires explicit description of the structural features that students must learn. Previous research has shown that the development of data representation usually begins with pictorial or concrete icons lacking a structured organization. Students may focus on one (possibly mathematically irrelevant) aspect of the data and not use relational or metarepresentational aspects (Watson & Fitzallen, 2010). This behaviour corresponds to our prestructural and emergent levels. At the structural level, students may draw icons in rows with the correct numbers of icons vertically aligned, so that frequencies can be compared using the lengths of the rows; prestructural students would probably not align the icons. The basic structural feature here is the principle of elementary graphical representation that equal numbers can be represented by equal lengths. In the subsequent development of formal categorical and continuous graphical representation, spatial structuring is necessary in visualising and organising representations along with relational understanding of the concepts of equal spacing, unitising, coordination of axes, and scale.

The present study investigated empirically the role of all these possible structural features in the development of graphical representation. By choosing highly able students as the subjects of the investigation, it was hoped to be able to observe in a relatively short time period what otherwise might be a lengthy development.

Method

Participants

Participants were 21 students in an academically selective Grade 1 class, all male, of an independent school in an Australian city. The students ranged in age from 6 years and 1 month to 7 years and 10 months (mean 6 years and 6 months) at the beginning of Grade 1, following their participation in PASMAP during Kindergarten. Students came from high socio-economic backgrounds and a range of cultural/ethnic groups.

Assessment

As measures of students' high ability, the researchers administered the Peabody Picture Vocabulary Test, 4^{th} edition (PPVT4) (Dunn & Dunn, 2007) and the Raven Coloured Progressive Matrices (RCPM) (Raven, 2004) in October of their Kindergarten year. The median score on each test was at the 95th percentile. Students were also administered two forms of the PASA interview at the beginning of Kindergarten and at the end of Grade 1 (Mulligan & Mitchelmore, in press). Broadly, the majority of students were classified as operating at the structural or advanced structural level at the Grade 1 interview.

Data Collection and Analysis

On the basis of assessment data and in consultation with the Grade 1 classroom teacher, students were divided into two learning groups: advanced (10 students) and less advanced (11). Each group was withdrawn from the regular class mathematics lesson for one hour per fortnight for four consecutive school terms (16 sessions) for the data modelling program. A suitable room was provided for this purpose. The classroom teacher was consulted about the planning and implementation of the program and was debriefed following each session. The lead researcher and an assistant led the sessions.

All student work samples, including their own written accounts of their activities, were collected and analysed along with observation and evaluation notes taken by the researchers during and after the learning sessions. Student data were collated in a student profile. Collection of video data was not permitted. Work samples were analysed for features of AMPS and subsequently coded for level of structural development based on the features revealed. Analysis of work samples employed an iterative process for comparing prior learning with new structural features (Lesh & Lehrer, 2000). Where appropriate, comparisons were made with data drawn from the students' PASA transcript.

Data Modelling Learning Sessions

An overall aim of the learning sessions was to develop a pedagogical approach to structuring data for young students prior to the use of technological tools. The program followed a series of tasks focused on several topics of interest that were relevant to the students' school learning program and incorporated a variety of opportunities for representing categorical and continuous data (*Pets in our class, Birthdays, Holiday destinations, Daily temperature, Melting ice, Growth of chickens*, and *Growth of onions*).

The sessions began with opportunities to discuss the topic and suggest ways of collecting and representing data. So that students could develop intuitive skills in visualising and sketching data, they were given minimal instructions on how to represent data and usually drew freehand on blank paper. Students' iterative refinement of their representations was important. After questioning of students individually about their representations from prior and present sessions, minimal scaffolds enabled them to improve the clarity and mathematical detail of their graphs. Students in the advanced group usually made two or three refinements per session; the other group made one. Students frequently compared representations and discussed similarities and differences and strategies for drawing improved graphs. Students justified their approaches to representing the data and explained the mathematical and scientific ideas inherent in the data.

Results

We provide examples of the development of AMPS within the graphical representations created for two topics, *Our pets* and *Daily temperature*. In the first example, students were given a template of a two-way table (with columns for type of pet, tally and number) to assist in recording data they collected from the class. They were then given a partially completed pictogram to complete. Figure 1 shows Mark's[2] initial attempt to represent the data without any scaffolding from the teacher. He self-corrected his one to one matching and added vertical and horizontal grid lines to structure the data. He explained that "there can only be one animal in each box if you want to count them".

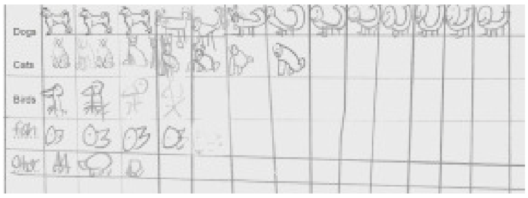

Figure 1. Mark's initial attempts to represent the data on *Our pets*.

Two weeks later we asked students to use 2 cm grid paper to record the data as a vertical pictogram, a task that most students found easy. We then asked them to reproduce their graph from memory. Figure 2a shows Mark's first attempt. Instead of different pictures, he used a single icon, a box. We may surmise that he was reproducing the boxes that contained the icons in his previous representation (Figure 1). He drew all his boxes roughly the same size and added numbers to represent the number of boxes to save the reader the effort of counting them. When asked to re-represent the data in a more accurate way, he re-drew the graph using crosses instead of boxes ("because they are easier to draw") and horizontal lines to align the crosses (Figure 2b).

Figure 2c depicts Mark's final attempt to represent the data as accurately as possible. He chose to use an even more efficient method, vertical lines, and coordinated the totals with the vertical scale. Note, however, that he did not space the vertical lines evenly. He shows relational thinking in that he returns to the table of data and shows the total number of pets as part of the graphical representation.

(a) (b) (c)

Figure 2. Mark's second set of graphs.

[2] Pseudonymns have been used to preserve anonymity.

In the second term of the program, students focussed on exploring representation of continuous data. The first investigation involved recording the daily temperature range in winter at various time intervals during the school day. Data for that day was retrieved from http://www.weatherzone.com.au/nsw/sydney. After interpreting the predicted minimum and maximum temperatures, the students then estimated the temperature at their morning session and predicted changes in temperature at 2- or 3-hour intervals for the rest of the day (some until midnight). Students used a considerable variety of representational strategies.

Edward's first graph (Figure 3a) used vertical lines instead of bars. On each line he drew a scale instead of numbers and appears to have attempted to use the same interval on all five lines. Edward also attempted to represent the temperature variation by outlining the pattern formed by the six scales—not, however, joining the tops of all of the lines.

Edward's second graph (Figure 3b) was a considerable advance. Not only did he use a common vertical scale but he recognised that it is not necessary to number every point on the scale and only labelled it at 5° intervals. It is now clear (as it was not in Figure 3a) that he was representing temperature by a point on a scale rather than an interval. Also, he correctly joined the tops of each line to show the temperature variation, but the horizontal spacing was still not correct.

(a) (b)

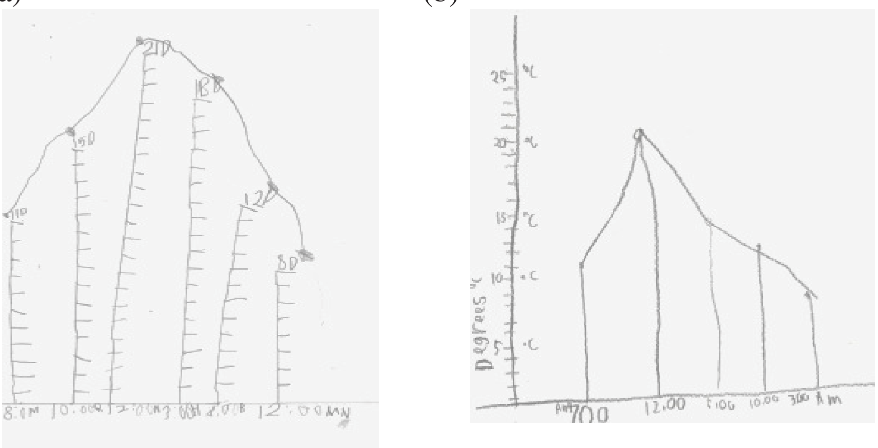

Figure 3. Edward's representations of the temperature variation.

Figure 4. Theo's temperature graph.

Figure 4 shows Theo's graph. He did not need to draw vertical lines to show each temperature point, showing that he could coordinate the vertical and horizontal axes. Notice that he has started his vertical axis at 10 instead of 0, presumably to save space, but he has again not scaled the horizontal axis correctly. Notice the students' difficulty of representing temperature variation between the measured times. They tend to 'join the dots' rather than constructing a smooth curve.

Discussion

Our data indicated there are common structural features that must be learned about frequency (and other categorical) charts and continuous graphs. The representation of a quantity is made by vertical bars/lines of various lengths (or points that represent their maximum). A common scale must be used on each bar or line, conveniently drawn on a vertical axis to the left. This scale defines points, not intervals. Not all values on the vertical axis need to be labelled, and the bars/lines/points should be appropriately spaced horizontally.

It was also clear that there are some differences between the two types of data and a danger of inculcating false ideas using bar charts. 'Interval scales' are natural for categorical data, but it is important to construct 'point scales'. The process of drawing bars/lines of a specific height should facilitate this transformation. Equal spacing on the horizontal axis is natural for categorical data (using abbreviated labels if necessary) but is not always appropriate for continuous data. 'Joining the dots' is natural for continuous data but usually inappropriate for categorical data.

In earlier PASA studies several links to aspects of structuring were found to be crucial. For example, equally spaced scales are used for spatially representing numbers (number lines), length (ruler) and time (analogue clock); skip counting and equal spacing in 2s or 5s are needed when deciding on scales to use. The vertical-horizontal structure of graphs is identical to the structure of a rectangular grid/array and includes ideas of congruence and colinearity. Students with a good understanding of the number lines and the rectangular grid structure may therefore be able to learn graphing skills more quickly than others.

An important observation in this study was students' ability to notice and then integrate all of the elements needed to create each type of graph. The use of meaningful real-life contexts, developed as a series of topics, meant that the concept of graph was built over time from experiencing different examples in different contexts. Allowing young students to create their own pictographs initially, without scale, was a basis for developing concepts of attribute, frequency and variation, to which they could later add scale.

Our data support the findings of English's (2012) analysis of representations in data modelling contexts. We found early representations were critical in assisting young children to extract meaning from their data (Konold & Higgins, 2003). We found a diverse range of icons, including structural features such as grid lines and symbols that reflected student's individual forms of representation. Further, the ability of these students to reflect upon and refine (re-represent) their representations was impressive. Students' lack of understanding of structure can remain hidden when they only face tasks where the structure is provided for them. In all the examples we have shown, drawing and refinement has been more revealing. Such tasks also challenge children to rehearse the structural features they have seen, and apparently understood at a superficial level, and to use these patterns to gain a deeper understanding of structure.

Conclusions and Implications

Our data enabled an exploratory analysis of the structural development of data modelling focused on graphing. We found that young, highly able students can develop over a short period of time critical developmental features that might be difficult to observe and describe systematically with less able students. While this sample indicated a risk of underestimating young children's capacity for metarepresentational competence, our findings do not permit generalisation. Nevertheless, we have been able to describe critical features that can be shaped into a pedagogical framework for data modelling. Implementation of curriculum priorities in statistical reasoning and associated teacher pedagogical learning may then be more easily achieved.

Acknowledgements

The research reported in this paper was supported by Australian Research Council Discovery Projects Grant No. DP110103586, *Transforming children's mathematical and scientific development: A longitudinal study*. The authors express thanks to Nathan Crevensten and Susannah Hudson, the Grade 1 teacher and the school community.

References

Dunn, L. M., & Dunn, D. M. (2007). *Peabody picture vocabulary test* (4th ed.). Bloomington, IN: Pearson Education.

English, L. D. (2010). Young children's early modelling with data. *Mathematics Education Research Journal, 22*(2), 24-47.

English, L. D. (2012). Data modelling with first-grade students. *Educational Studies in Mathematics Education*, *81*, 15-30.

Konold, C., & Higgins, T. L. (2003). Researching about data. In J. Kilpatrick, W. G. Martin, & D. Schifter, D. (Eds.), *A research companion to principles and standards for school mathematics*. Reston, VA: National Council of Teachers of Mathematics.

Lesh, R., & Lehrer, R. (2000). Iterative refinement cycles for videotape analyses of conceptual change. In R. Lesh & A. Kelly (Eds.), *Research design in mathematics and science education* (pp. 665-708). Hillsdale, NJ: Lawrence Erlbaum Associates.

Lehrer, R., & Schauble, D. (2005). Developing modeling and argument in elementary grades. In T. Romberg, T Carpenter, & F. Dremock (Eds.), *Understanding mathematics and science matters* (pp. 29-53). NJ: Erlbaum.

Mulligan, J. T., & Mitchelmore, M. C. (in press). *The Pattern and Structure Assessment—Early mathematics.* Melbourne, VIC: ACER.

Mulligan, J. T., & Mitchelmore, M. C. (2009). Awareness of pattern and structure in early mathematical development. *Mathematics Education Research Journal, 21*(2), 33-49.

Mulligan, J. T., English, L. D., Mitchelmore, M. C., & Crevensten, N. (in press). Reconceptualising early mathematics learning: The fundamental role of pattern and structure. In L. D. English & J. T. Mulligan (Eds.), *Reconceptualizing early mathematics learning*. New York: Springer.

Papic, M., Mulligan, J. T., & Mitchelmore, M. C. (2011). Assessing the development of preschoolers' mathematical patterning. *Journal for Research in Mathematics Education,* 42(3), 237–268.

Raven, J. (2004). *Coloured progressive matrices and Crichton vocabulary scale—Manual.* London: Pearson.

Watson, J., & Fitzallen, N. (2010). *The development of graph understanding in the mathematics curriculum: Report for the New South Wales Department of Education and Training.* Sydney: NSW Department of Education &Training.

Watson, J. M., & Kelly, B. A. (2005). The winds are variable: Student intuitions about variation. *School Science and Mathematics, 105*, 252-269.