GETTING ORGANISED: THE ROLE OF DATA ORGANISATION IN STUDENTS' REPRESENTATION OF NUMERICAL DATA

Steven Nisbet
Griffith University

This paper explores the role of organising data in representing numerical data in an organised way – a skill which does not come naturally to most students. Grade 6 children were given a series of lessons on organising data and then representing the data graphically. It was found that after instruction, more students were able to reorganise the data and produce an organised graph.

INTRODUCTION

This paper explores the effect of instruction on Grade six students' organisation and representation of numerical data. Previous studies (Nisbet, 2001, 2002, 2003 in press) have shown that while most students find it easy to re-organise and represent categorical data (e.g., eye colour), fewer students are able to do this with numerical data (e.g., number of books read).

By the end of Grade 6, Queensland students are expected to be able represent the data in tables, bar graphs, line graphs, circle graphs, histograms, etc. (Department of Education, 1987). The Australian Numeracy Benchmarks (Curriculum Corporation, 2000) state that the ability to organise, summarise, and display information graphically is essential for primary school students (i.e. Grades 1 to 7). Similarly, the National Council of Teachers of Mathematics Standards (NCTM, 2000) highlights the need for students at all school levels to organise and represent data.

Research into students' representation of data has included the development of a statistical-thinking framework by Jones, Thornton, Langrall, Perry, & Putt (2000). One construct in the framework - *Representing Data* - incorporates making representations that exhibit different organisations of the data. Four levels of thinking are proposed for the construct. Firstly, a Level 1 student produces an idiosyncratic display that does not represent the data in a valid way. A Level 2 student produces a valid display that represents the data but does not attempt to reorganise the data. A Level 3 student produces a display that not only represents the data but also shows some attempt to reorganise the data. A Level 4 student produces multiple valid displays, some of which reorganise the data.

Other studies provide further background for this study. Lehrer and Schauble (2000) investigated the process of data organisation with children in grades 1, 2, 4 and 5. Their results suggest that children at higher grades use more sophisticated strategies for organising data than those in lower grades. Nisbet, Jones, Thornton, Langrall, & Mooney (2003, in press) analysed Grades 1 to 5 students' representations of categorical and numerical data, and found that numerical data was more difficult for children to reorganise and represent than categorical data. Whereas 60% of the children were able to reorganise categorical data, only 20% could reorganise numerical data. Another study (Nisbet, 2001) found that many teacher-education students had similar difficulties organising numerical data. All produced an organised graph from categorical data, but only 19% produced an organised graph from numerical data, with the majority of students merely drawing separate bars for individual pieces of data without reorganising the data into numerical categories.

Why do many students at all levels find it difficult to reorganise numerical data? It could be the verbal nature of the categories (e.g. blue eyes, brown eyes) make them more obvious compared to numerical data (e.g. 2 fish, 3 fish, etc.). It may be that the perceived need for reorganisation depends on the size of the data set – the larger the data set, the greater perceived need to organise the data. To test this hypothesis, Nisbet (2002b) asked students in Grade 9 and 11 to draw graphs of two sets of numerical data, one with 10 pieces of data, and one set with 30 pieces of data. With the smaller set, most students drew bar graphs showing no reorganisation of the data – just individual bars. However, with the larger set, more students reorganised the data according to frequencies of scores and then drew an organised representation. Those students having difficulty in reorganising the data were given prompts drawing attention to the frequency of the numerical values in the raw data. For the Grade 11 students, the ability to organise the data without prompting was greater for those of high mathematics ability. However, there was no similar ability effect for Grade 9 students. In another study (Nisbet, 2002a), the majority of Grade 7 students needed prompting to reorganise the larger set. However, the less mathematically able students required more prompts than their more able counterparts.

If students find the task of reorganising and representing large sets of numerical data difficult, and if these skills are not only desirable but also necessary in terms of the core curriculum, then ways need to be found to assist students to acquire these skills in a meaningful way. Pertinent research questions such as the following therefore arise: (i) To what extent does instruction assist students to learn skills in reorganising and representing numerical data? (ii) What teaching/learning activities are effective in assisting the students learn these skills? (iii) What benefit would be obtained by employing a *dynamic* approach (Russell & Friel, 1989) in which students investigate issues of interest to them and collect, analyse and represent their own data, in contrast to a *mechanistic* approach (Ernest, 1989) in which students are taught separate skills in a rule-based way, and analyse second-hand data; and (iv) What is the role of mathematics achievement in reorganising and representing data.

METHOD

Participants

The participants were 50 children in a double Grade 6 class at a government school in Brisbane, Australia. Their mathematics ability ratings (as judged by their teachers) were as follows: A+ (very high): 2 students; A (high): 5 students; B (good): 17 students; C (average): 17 students; and D (below average): 9 students. Six target students (across ability levels) were chosen for follow-up interviews. The study was conducted in the last weeks of the school year. The children had completed almost six years of schooling, and (according to the syllabus) should have had lessons each year on collecting and organising data, and constructing graphs to represent data.

Design

The study took the form of a teaching experiment (Cobb, 1999) in which students encountered a series of learning activities, and their performance was studied. The researcher worked with the two teachers to administer a pre-test, conduct the three lessons, and administer a post-test. These lessons covered methods of organising and representing data, and utilised data the students collected about themselves on topics of interest. The pre-test and post-test results were subsequently analysed.

Data collected

At the outset, the students were given a pre-test on organizing and representing data. The tasks were printed on paper for each student (Figure 1).

The following are data on the number of books read by 25 students in a term.

2	3	6	3	4	2	0	5	5	3	1	2	4	4
4	3	3	2	5	5	6	1	2	3	3			

Task 1: Organize the data in some way that makes sense to you.

Task 2: Represent the data in any way you like.

Figure 1: Pre-test task

The students' responses were assessed in terms of the framework (Jones et al, 2000) and the six target students were interviewed along the following lines: (i) Please explain how you organized the data, and why it makes sense to you. (ii) Please explain how you represented the data and why you did it that way.

Two days after the last of the three lessons, the class was given a post-test on the topic (almost identical with the pre-test) and the six target students were interviewed about their responses. The post-test items are showed in Figure 3, and the interview questions were the same as for the follow-up to the pre-test (Figure 2).

The following are data on the number of CDs bought by 25 students in a year.

3	4	6	3	4	2	0	5	5	7	1	2	4	4
4	3	3	2	5	5	6	1	2	3	3			

Task 1: Organize the data in some way that makes sense to you.

Task 2: Represent the data in any way you like.

Figure 2: Post-test task

Series of lessons

The lessons focused on the processes involved in collecting, organising and representing data, with particular attention to organising data. The *first lesson* focused on organizing some raw data on a topical issue presented to the students. (In subsequent lessons students collected data about themselves). After a brief discussion on healthy foods and take-away food, the students were presented data on take-away meals eaten by children in a hypothetical class (Figure 4).

The researcher demonstrated ways of organizing data (rank order, line plot, and tally table), and the students worked in pairs to reorganise and represent the data. In the *second lesson* the students collected, organised and represented data in a 15-item survey about their physical measurements, eating habits, plus estimates of time, length, distance, and number (Figure 5). This lesson lasted $1\frac{1}{2}$ hours, and the students (to the teachers' amazement) stayed on task for most of the time.

Number of take-away meals eaten by children in the last month:

2	**4**	**2**	**7**	**4**	**0**	**4**	**3**	**5**	**1**
5	**8**	**1**	**7**	**5**	**1**	**3**	**4**	**6**	**5**

Figure 3: Data for the lesson on organizing data

The *third lesson* included another survey, this time about topics suggested by the students, e.g., sport music, movies, DVDs, pets, families, careers, and pocket money (Figure 4). Students collected the data, then, in pairs, organised the data with line plots and tally plots, then drew organised graphs.

Sections	Items
Physical measure-ments	Your handspan in cm (fingers stretched apart) Your height in cm Your vertical reach in cm (one arm up) Your horizontal reach in cm (two arms stretched out)
Eating habits	How many times have you bought food from the tuckshop in the last 4 weeks? How many times have you eaten a bowl of cereal in the last 7 days? How many times have you eaten rice in the last 7 days? How many times have you drunk a can of soft drink in the last 7 days? How many times have you drunk a glass of milk in the last 7 days? How many pieces of fruit have you eaten in the last 7 days?
Estimates	Estimate of time interval (seconds) Estimate of length of rope (metres) Estimate of distance from school to city (km) Estimate of number of dice in a jar (to nearest 10)
Other	Time it takes you to travel to school (in minutes)

Figure 4: Items in the class survey for Lesson 2

The planning of instructional activities involved some logistical challenges. The first was motivating a double class, and keeping them interested and on-task. The keys to meeting this challenge were threefold: (i) ensuring the topics were interesting and relevant; (ii) having adequate space for whole-class and small-group activity; and (iii) having three teachers (two teachers plus researcher) to monitor the students. Re the first point, the topics selected for study were related to the students' world. As for space, the double classroom had generous dimensions, allowing a carpeted area for all children to sit close to the teacher, plus plenty of space for table clusters.

The second challenge was to collect data from the double class efficiently, and distributing the data to the pairs of students to analyse. The first technique devised was to print a survey sheet with 15 items listed twice (in parallel columns, as in Figure 5), and to have the students fill in their responses twice, producing 30 responses each. The students then cut the sheet into two halves and each half into 15 pieces. They then deposited the pieces into numbered plastic ice-cream containers. Each pair of students, therefore, received a large set of data to analyse.

Prior to filling in the sheets the students needed to measure their heights, handspans, etc., requiring appropriate measuring equipment to be available around the room. The second technique was to have a data collection sheet for each survey item, and to get the children to complete them in turn, and pass them around a single class.

1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15

Figure 5: Survey sheet format for Lesson 2

RESULTS AND DISCUSSION

The results show an improvement in the students' reorganisation and representation of numerical data. A comparison of pre-test and post-test scores reveals an increase in the number of students who were able to draw an organised graph – from 19 out of 50 students (38%) in the pre-test to 47 out of 50 (94%) in the post-test.

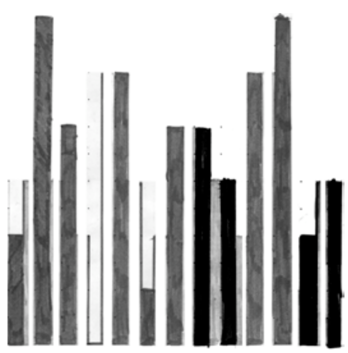

Fig. 6: Graph showing no reorganisation (pre-test)

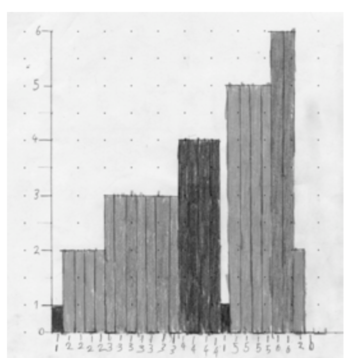

Fig. 7: Graph organised by magnitude (pre-test)

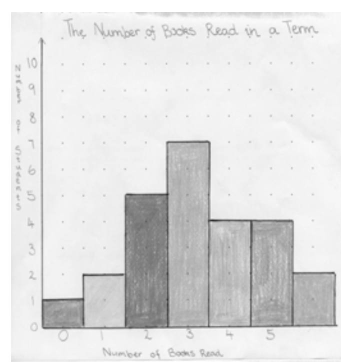

Fig. 8: Graph organised by numerical categories (pre-test)

In the *pre-test*, 31 students (62%) drew individual bars showing no re-organisation of the data (Figure 6). One student's graph also had 25 separate bars, but was organised by magnitude, (mostly) (Figure 7). The other 18 students drew graphs organised by categories, showing frequency of students versus number of books read (Figure 8).

Methods of reorganising the data varied among the students. Of the 18 students who drew graphs organised in categories, six made a tally table first. Six students just jotted down how many people had read the numbers of books. Another student had circled or highlighted the numbers of books in different colours to help her reorganise the data – the 1s circled in black, the 2s highlighted in green, etc. The other students apparently just counted the frequency of occurrence of the scores without noting the frequencies.

The *interviews after the pre-test* revealed that the two A-level students had in fact counted the frequency of each score, and drawn their graphs from that information. One made a tally table. One B-level student thought about making a table but did not proceed. One of the D students (the one who colour coded the data) explained that the colours made it easy for her to count the frequencies. The other D student interviewed saw that 25 bars from 25 pieces of data would not fit the 13 divisions on the horizontal axis of the graph paper, so he split the divisions to fit all the bars in.

Regarding the *post-test graphs,* the majority (47 out of 50) showed a reorganisation of the data (Level 3). One showed bars in order of magnitude (similar to Figure 8). and 46 showed frequency versus numbers of CDs (similar to Figure 9). The other 3 graphs showed no reorganisation (Level 2) (similar to Figure 7), despite the fact that these students initially reorganised the data in a tally table or line plot. One of the teachers observed that one student had drawn the unorganised graph first, then reorganised the data in a tally table *and* a line plot afterwards.

Further in the *post-test graphs,* the majority of students demonstrated the ability to reorganise the data prior to drawing their graphs; 48 out of 50 made a tally table, and of these, 34 drew a line plot also. The other two drew a line plot only. It is surprising that three students went to the effort of reorganising the data in a tally table, but did not translate this reorganised data into an organised graph. Two of these students were rated by the teachers as B students, and the other as a D student. There is room for further research here on what prevents such students from drawing a representation of the reorganised data.

The *post-test interviews* revealed that students had no difficulty in reorganising data with either a table or line plot. These are obviously appropriate tools for reorganising data for Grade 6 children. The students understood the need to scan the raw data to note the lowest and highest values for the tally table, and readily adopted this as part of the procedure to reorganise the data. Also, the students revealed that they could interpret their own graphs i.e., they could explain what each bar meant. However, one D-level student reversed the meanings of the axes and confused the numbers of CDs with the number of students. One B-level student interviewed was one of three who had made a line plot to organise the data, but then drew a graph with individual bars. When asked why she drew it that way, she referred back to the original raw data, not her own reorganised data. She remarked, "The first person bought 2 CDs, the second bought 3 CDs, …". It appears that for some students the links between the reorganised data and their graphical representation need to be made very explicit through student-teacher dialogue concerning the reorganised data and the graph.

There was no significant *effect of mathematics rating* on the level of graphs produced in the pre-test, nor the post-test, nor on the improvement in levels. Although the two A+ students produced graphs showing reorganisation of the data (Level 3) in the pre-test, so also did four of the nine D students. In the post-test, two of the three students who did not draw organised graphs were B students; the other was a D student. This result could imply that (i) the teachers' ratings may not be accurate; (ii) the skills required to reorganise and represent numerical data are fairly specific and not entirely coincident with general mathematics skills. Thus, it would be worthwhile to obtain a better measure of mathematics ability – a standardised measure with spatial and logical components as well as numerical components.

There was one unanticipated observation arising from this study – an association of pre-test results with *music experience*. When the researcher arrived for the pre-test, the teachers informed him that 17 of the students – members of the school orchestra – would be away at a recording session. Nevertheless, the teachers indicated that they would administer the pre-test to them later in the day. Hence the researcher had the pre-test papers in two separate bundles, and was able to code them according to orchestra membership. The results revealed that 9 of the 17 orchestra members (53%) had reorganised the data before drawing their graphs (Level 3), whereas 9 of the 33 non-members (27%) had reorganised the data before drawing their graphs (Level 3). Although a chi-square test showed no significant relationship between orchestra membership and level of pre-test graph, with more sensitive measures of music ability and handling data. the relationship could be explored further.

The possibility of an association with music ability is interesting in view of previous studies which show links between mathematics and music ability (Nisbet & Bain, 2000; Rauscher & Shaw, 1993). However the result here should not be taken any further at this point given that the teachers indicated that the orchestra members were "some of their brightest students". An analysis of variance revealed a significant association between teachers' ratings mathematics ability and orchestra membership ($F = 10.344$, $p = .002$). It appears that the school orchestra attracts students of high mathematics ability. This result offers another opportunity for further research.

In summary, the most important finding of this study was the increase in the number of students who were able to reorganise numerical data and validly represent them. The study demonstrated that data-reorganising skills can be taught to Grade 6 students meaningfully, that such skills assists them to draw valid organised graphs, and that learning activities that involve students in collecting and analysing data about themselves are effective in sustaining interest and engagement. The study has also identified avenues for further research relating to the transition from reorganised data to their representation, utilising more sensitive measures of data handling, mathematics ability and music ability.

References

Cobb, P. (1999). Individual and collective mathematical development: The case of statistical analysis. *Mathematical Thinking and Learning, 1,* 5-43.

Curriculum Corporation (2000). *Australian Numeracy Benchmarks*. Melbourne, Vic.: Author.

Department of Education, Queensland (1987). Years 1 to 10 Mathematics Syllabus, Year Overviews. Brisbane: Curriculum Services Branch

Ernest, P. (1989). The impact of beliefs on the teaching of mathematics. In P. Ernest (Ed.), *Mathematics Teaching: The state of the art*. New York: Falmer.

Jones, G., Thornton, C., Langrall, C., Mooney, E., Perry, B., & Putt, I. (2000). A framework for characterizing students' statistical thinking. *Mathematical Thinking & Learning, 2*, 269-307.

Lehrer, R., & Schauble, L. (2000). Inventing data structures for representational purposes: Elementary students' classification models. *Mathematical Thinking & Learning, 2*, 51-74.

National Council of Teachers of Mathematics (2000). *Principles and standards for school mathematics*. Reston, VA: Author.

Nisbet, S. (2002a). Year 7 students' representation of numerical data: the influence of sample size. In B. Barton, K. Irwin, M. Pfannkuch, & M. Thomas (Eds.) *Mathematics Education in the South Pacific, Proceedings of the 25th Annual Conference of the Mathematics Education Research Group of Australasia,* Auckland, NZ, July 7-10, 2002. (pp. 520-527). MERGA.

Nisbet, S. (2002b). Representing numerical data: the influence of sample size. In A. Cockburn & E. Nardi (Eds.) *Proceedings of the 26th Annual Conference of the International Group for the Psychology of Mathematics Education*. PME 26, Norwich, UK. July 21-26, 2002. Vol. 3, pp. 417-424. Norwich, UK: University of East Anglia.

Nisbet, S. (2001). Representing categorical and numerical data. In J.Bobis, M.Mitchelmore, B.Perry (Eds.) *Proceedings of the 24th Annual Conference of the Mathematics Education Research Group of Australasia*, Sydney, NSW, 1-4 July 2001. Sydney: MERGA.

Nisbet, S. & Bain, J. (2000). Listen to the graph: Children's matching of melodies with their visual representations. In T. Nakahara & M. Koyama (Eds.) *Proceedings of the 24th Annual Conference of the International Group for the Psychology of Mathematics Education*. PME 24, Hiroshima, Japan. July 23-27, 2000. Vol. 4, pp. 49-56.

Nisbet, S., Jones, G., Thornton, C., Langrall, C., & Mooney, E. (in press). Children's Representation & Organisation of Data. Accepted for *Mathematics Education Research Journal 15* (1).

Rauscher, F.H., Shaw, G.L., Ky, K.N. (1993). Music and spatial task performance, *Nature*. 365:611.

Russell, S., & Friel, S. (1989). Collecting and analyzing real data in the elementary school classroom. In P. Trafton & A. Schulte (Eds.) *New Directions of Elementary School Mathematics, 1989 Yearbook*. Reston, Va: National Council of Teachers of Mathematics.

Wood, D., Bruner, J., Ross, S. (1976). The role of tutoring in problem solving. *British Journal of Psychology, 66,* 181-191.