Graphs: Communication lines to students?

Jonathan B. Moritz and Jane M. Watson
University of Tasmania

Graph comprehension is considered a basic skill in the curriculum, and essential for statistical literacy in an information society. How do students interpret a graph in an authentic context? Are misleading features apparent? Responses to questions about a graph-based advertisement suggest that students commonly did not appreciate scaling difficulties, relate a graph as relevant in the context of a standard interpretation task, or apply numeracy skills for calculations based on data in graphical representations.

According to a proverb, a picture tells a thousand words, though just which thousand words may depend on how the picture is "read." In our information age, displaying and interpreting graphical information is fundamental to being statistically literate. The increasing presence and variety of graphs associated with claims in the media raises the potential for misleading the public, with the claim of authority, "Statistics show...". Some misrepresentations may be inadvertent errors, such as producing a pie chart accounting for 128% of whole. Others may use purposely misleading representations, such as cutting off the vertical axis of a bar chart to make differences appear relatively greater than they in fact are. As advertisers use more graphs to convince potential customers of the merits of their products, readers need to recognise misleading representations and protest their use.

Recent curriculum documents (Australian Educational Council [AEC], 1991, 1994; Ministry of Education, 1992) have emphasised data representation as a part of the data handling curriculum relevant to mathematics, science and social science. In *A National Statement on Mathematics for Australian Schools* (AEC, 1991), statement B5 suggests students should "represent, interpret and report on data in order to answer questions posed by themselves and others," including examples such as

- Represent data in tables and graphs and compare different representations of the same data, considering how well they communicate the information (e.g. correct, clear, misleading).
- Discuss and interpret information presented in graphs and tables found in newspapers, magazines and text materials. (p. 172)

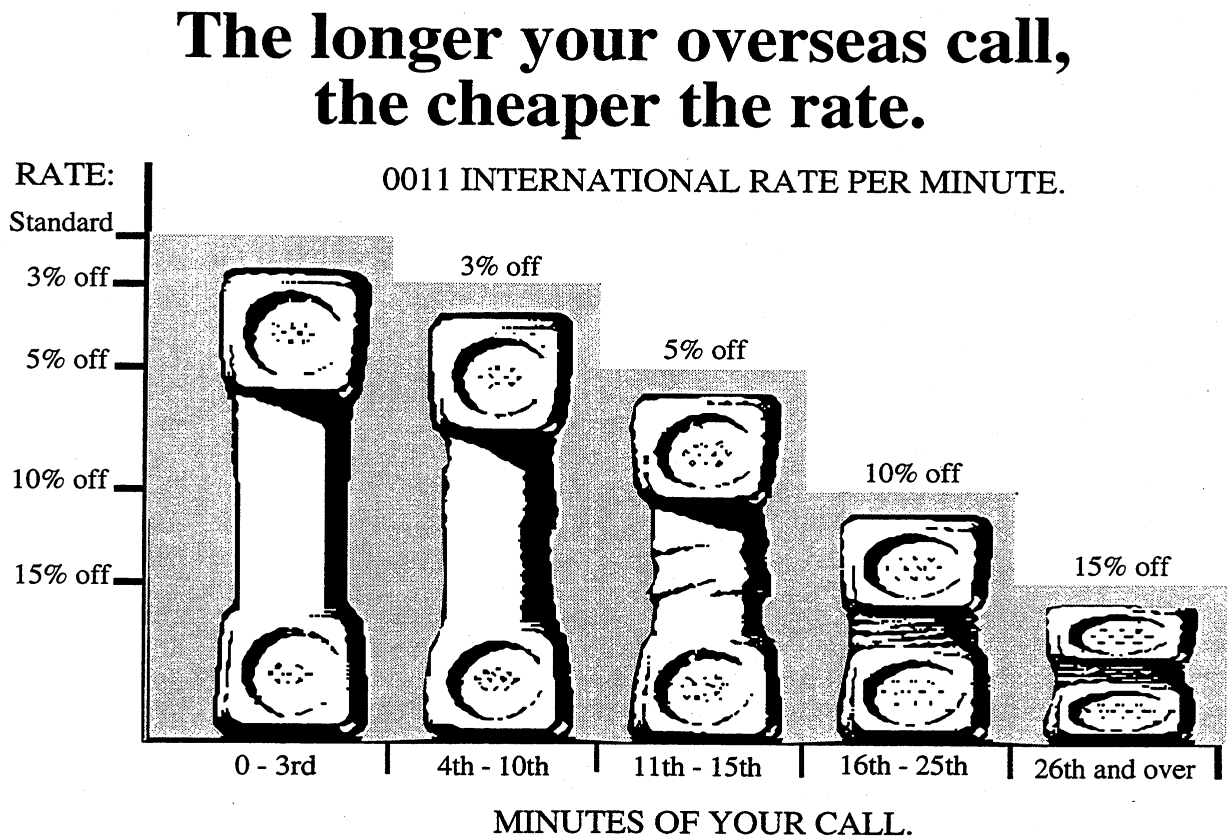
Similarly, in *Mathematics - A curriculum profile for Australian schools* (AEC, 1994, p. 93), element 5.27 includes the example requirement to

- Interpret bar graphs and histograms for grouped data, including where the scales on the axes must be read between calibrations.

Research involving students' comprehension and interpretation of statistical graphs appears to have been less than prolific. With respect to young children's understanding and use of graphs, Pereira-Mendoza and his co-workers have identified a number of areas of possible misunderstanding: interpreting scale, distinguishing missing values from zero value data, mis-using topic knowledge when predicting, making unfounded assumptions about patterns, finding intermediate values, and manipulating bar graphs (Pereira-Mendoza & Mellor, 1991; Pereira-Mendoza, Watson & Moritz, 1995; Watson & Pereira-Mendoza, 1996).

Research by Curcio (1987) with students in grades 4 and 7 led her to suggest three levels of questions teachers should ask in relation to graph comprehension: (1) reading the data, (2) reading between the data, and (3) reading beyond the data (Curcio, 1989). At level (1), the "reader simply 'lifts' the facts explicitly stated in the graph, or the information found in the graph title and axis labels, directly from the graph" with no interpretation (p. 5). At level (2), the reader "requires the ability to compare quantities ... and the use of other mathematical concepts and skills ... that allow the reader to combine and integrate data" (p. 6). At level (3), the reader must use background knowledge "to predict or infer from the data" (p. 6). Similar distinctions have been used to set levels of questions about a bar chart (Watson & Moritz, 1996a) and a pictograph (Moritz & Watson, 1997) involving (1) literal reading of data values, (2) calculations using intermediate results to reach a conclusion, and (3) predictions based on the given data as well as topic knowledge. While these levels are reasonably distinct, it was found that some level (2) questions, such as the highest value, were easier for students than level (1) questions, because they did not require reading the scale of the graph for the data value to answer the question.

More broadly in statistics education, Watson (in press) has suggested a hierarchy of statistical literacy which involves (i) having knowledge of basic statistical terms, (ii) recognising, interpreting and using these in applied contexts, and (iii) being able to question unrealistic claims made by the media or others. While applied more widely to concepts relevant to items presented in the social contexts of the media, it has similarities to the requirements of Curcio's levels of graph comprehension. Certainly reading beyond the data is necessary for questioning claims associated with graphs which are made in the media. Examples of the applications of the statistical literacy hierarchy include explaining the nature of a pie chart which sums to 128% (Watson, in press), and drawing a graph to represent a claimed cause-effect relationship between automobile usage and heart deaths (Watson & Moritz, 1996b).



(a) Explain the meaning of this graph.

(b) Is there anything unusual about it?

Suppose the standard rate is $1.00 for 1 minute.
You have already talked for 30 minutes.
(c) How much would the next 10 minutes cost?

(d) How much did the first 30 minutes of the phone call cost?

Figure 1. Graphing item from media survey about Chance and Data

This report aims to explore the ways students interpret a graph in the context of a potentially misleading newspaper advertisement in the light of the hierarchy for statistical literacy above. The item chosen is shown in Figure 1. The graph presents an interesting challenge to statistical literacy and may cause the observer to question the motives of the company which produced it. Both the vertical and horizontal scales are non-linear and the vertical axis is truncated with a somewhat unusual feature of values appearing to decrease with height. Many basic numeracy skills are required to interpret the graph and the implications of the information it contains. In Part (a), students are expected to describe the meaning of the graph to indicate basic understanding of the representation in context. Part (b) asks students to identify any unusual aspects of the graph, which allows students to critique the advertisement at the third level of the statistical hierarchy, by questioning the representation. Parts (c) and (d) require numeracy skills involved in reading values off the graph and then interpreting them in context with calculations. The calculations in Parts (c) and (d) are the sort which might be expected of members of the community who want to know of the benefits offered by the company. Particular attention hence is given to the ways students employ numeracy skills to interpret the data represented and to the ways students evaluate how well the graph and the entire advertisement represent and communicate the information.

Method

The item in Figure 1 is from a media survey which was part of a study of student understanding of concepts in the chance and data curriculum (Watson, 1994b; Watson, Collis, & Moritz, 1994). The survey was administered during 45 minutes of class time to over 1800 students in Grades 6, 8, 9 and 11 at Tasmanian government schools during 1993 and 1995 (sample sizes for groups are given with the results). This report compares performances at different grade levels, and considers change in performance for students who answered the item in both years. Analysis of individual categories of response will use a diagrammatic procedure for mapping responses (Watson, 1994a) to illustrate the structure of responses and analyse common errors. This procedure highlights the data used in problem solving, the concepts and processes involved, and how these are combined to produce intermediate and final responses.

Results

Part (a) Describing the graph

Responses about the meaning of the graph often involved minimal re-wording of the advertisement title, corresponding to Curcio's (1989) literal "reading the data." Students often lifted the title from the advertisement to explain the meaning, or re-worded it slightly, often preferring "cost" to "rate." In this respect, it is not evident whether these students understood this distinction.

> The longer your overseas call, the cheaper the rate. [Grade 9]
>
> The longer you are on the phone, the less you pay. [Grade 9]

Some responses did express in the students' own words a way of interpreting the graph. The first response below appears to initially use the title to make a simplified statement about cheaper calls, but goes on to describe the percentage discount, without clarity as to whether this discount applies to the final time interval or to the entire call. The second response makes clear the discount rate applies for different time intervals.

> The longer you spend talking on the phone to an overseas person the cheaper your call becomes, e.g., if you spend 26 minutes on the phone you get 15% off the cost of your call. [Grade 9]
>
> Standard rate gets cheaper as the minutes pass, while talking on the phone ie. first 3 mins. costs standard rate, next 1-6 mins costs 3% less than standard, next 1-5 mins costs 5% less than standard, next 1-10 mins costs 10% less than standard. [Grade 11]

The following responses illustrate a variety of ways that students misinterpreted the graph, showing they did not appreciate the variables written as the two axis labels.

> It tells you how many calls you make. [Grade 6]
>
> It shows you when to ring, e.g., ring 4-10 October and it's 3% off your normal phone call. [Grade 9]

These disturbing responses may give an indication why many were unable to make sense of Parts (c) and (d) discussed below. Overall less than 10% of responses could be confidently categorised as going beyond the title to state the relatively complex meaning of the graph.

Part (b) Questioning unusual features of the graph

Most students responded that they considered nothing unusual about the graph. Some responses considered the graph unusual based on what students think a phone or a graph should look like.

> There's no way you can do that to a phone. In the last phone, the ear part of the phone wouldn't reach when you're talking in it. [Grade 9]
>
> They used phone instead of a line graph. [Grade 6]

Some responses had difficulty comprehending that decreasing rate and increasing cost are compatible. The first two responses below questioned the claim, based on the students' misunderstanding that it referred to reduced cost rather than reduced rate. The final two responses appear to understand the distinction, and then proceed to question the clarity of the advertisement and consider how it might be misleading for some.

> Yes, the longer you call it costs extra for lines etc. Why does it get cheaper? [Grade 8]
>
> It should cost more the longer you talk but it doesn't. And if you talk for a great amount of time it might end up being a 100% off. [Grade 9]
>
> Yes, at first glance you'd think you could stay on there all day and not pay anything, but you can't. [Grade 9]
>
> The presentation is a bit silly (the phones). The untrained eye might think their call actually got cheaper! [Grade 9]

Some students were able to comment on the unusual nature of the representation with respect to the reversed scale, which appeared to be recognised as unusual in contrast to standard expectations, almost assuming patterns in graphs.

> The highest rate of discount is at the bottom - the start is higher and it declines when graphs usually incline. [Grade 9]
>
> It's not drawn to scale. It's kind of back to front. You'd think the big phone would represent big saving. [Grade 9]

Other responses commented on the non-linear aspect of the scales, acknowledging that this gives a visual misrepresentation which might mislead people.

> The space from 3% - 5% is the same as the space form 10% - 15%. [Grade 8]
>
> On the graph, the 15% mark is around 1/4 of the original price. Once you reach 26 minutes the charge stays fixed. [Grade 9]
>
> The size of the phone handle is not representative of what percent you get off. [Grade 9]
>
> The prices only go down a fraction, and not alot, like most people would think it was. [Grade 9]

Table 1 shows the distribution of categories of response by grade level. Less than 10% of students commented on the unusual scale or misrepresentation of the graph. The high proportions of students offering no response, or a response that nothing was unusual about the graph, suggest that while between 25% and 49% of students (depending on grade) appreciate that generally this graph may be unusual, few students have reached level (iii) of the statistical hierarchy in relation to this task.

Table 1
Percentage of Responses to Part (b) by Response Category and by Grade

Response Category	Grade 6 1993	Grade 6 1995	Grade 8 1995	Grade 9 1993	Grade 9 1995	Grade 11 1995
Scale/Reverse	1	2	1	9	6	8
Phone/Cost	24	32	31	40	37	35
No/Yes	58	52	50	37	43	35
NR	17	14	18	15	15	21
N	*286*	*216*	*295*	*329*	*279*	*179*

Parts (c) and (d) Applied use of the graph with calculations

Responses to Parts (c) and (d) varied greatly in structural complexity. The mappings of responses shown in Figures 2 to 9 illustrate paths by which intermediate and final responses may have been reached by students, and assist in explaining errors in calculation or interpretation. For Part (c), the simplest structural form of response involved using the cues "10 minutes" and "rate of $1/min" to yield the response of "$10 cost," as shown in Figure 2. A similar response which also uses the cue "30 minutes" results in a response "$40 cost," is shown in Figure 3. The most significant feature of these responses is that they show no evidence that the graph plays any role in the students' reasoning for the task at hand. The majority of Grade 6 students responded in one of these ways, as is seen in Table 2. The percentage decreased with increasing grade level.

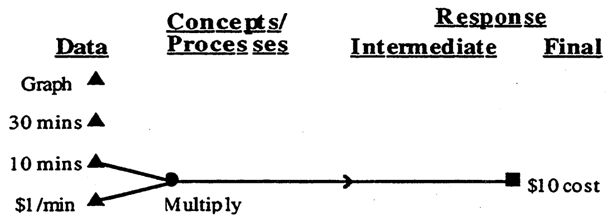


Figure 2. Map of $10 response to Part (c)

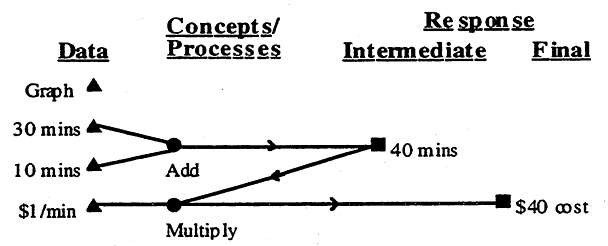


Figure 3. Map of $40 response to Part (c)

Of those who did employ the graph in their response, some students used the graph along with a single other cue to select a value to read off the graph. Those who used the cue "10 minutes" responded with an inappropriate response of "3% off," as shown in Figure 4. Other students responded "15% off" by using the cue "30 minutes" to select the appropriate value to read off the graph, as shown in Figure 5. These students, however, did not go on to use this to calculate a correct response of the cost for another 10 minutes. The category of responses related to "15%" included those who went on to use "15% off" to give responses of 15c, 85c, or $1.50. These inappropriate or incomplete attempts together represent between 3% and 12% of responses for various grades in Table 2.

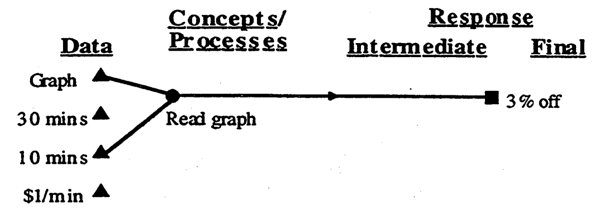


Figure 4. Map of 3% response to Part (c)

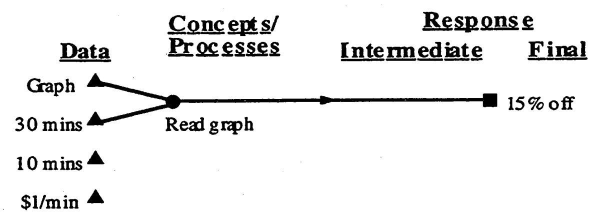


Figure 5. Map of 15% response to Part (c)

The correct answer of $8.50 can be obtained in several ways, two of which are shown in Figures 6 and 7. The need for greater structural complexity to find the correct answer is seen in these figures. The percentage correct increases with grade (Table 2). Other responses which did not fall into one of the categories above were either considered as non-response (NR), or placed in the "Other" category. Included in the latter were responses of $7.50, although it is not clear how students achieved this response, i.e., whether they misinterpreted the graph or made a calculation error.

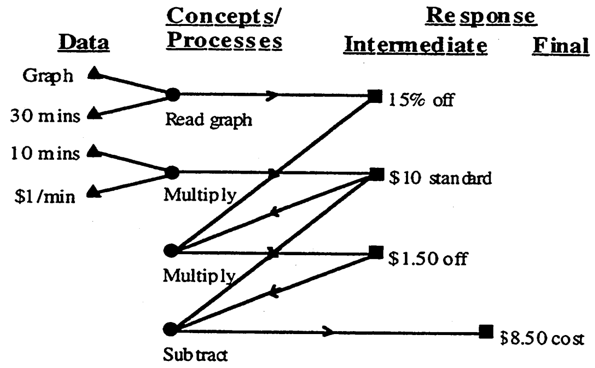


Figure 6. Map of $8.50 response to Part (c)

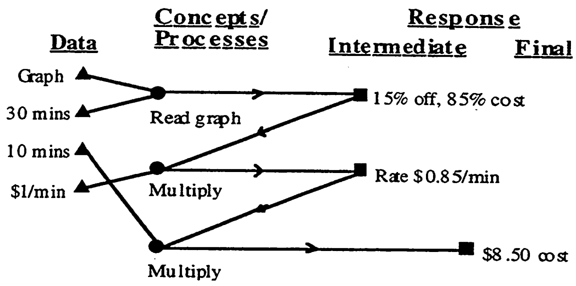


Figure 7. Map of $8.50 response to Part (c)

Table 2
Percentage of Responses to Part (c) by Response Category and by Grade

Response Category	Grade 6 1993	Grade 6 1995	Grade 8 1995	Grade 9 1993	Grade 9 1995	Grade 11 1995
$8.50	4	3	10	21	14	28
15%/$1.50/85c	2	2	3	6	9	11
3%...	4	1	3	4	1	1
$10/$40	64	64	43	21	25	16
Other	14	14	15	24	20	18
NR	12	15	26	25	30	27
N	*286*	*216*	*295*	*329*	*279*	*179*

For Part (d), similar categories of response were found - those which ignored the graph to respond $30, and those which used the graph to get a percentage value, some of which went on to use to calculate a cost, either based on a 15% discount for the entire call yielding $25.50, or based on the correct interval stages of discounts to get a correct result of $27.79. Figures 8 and 9 show this is a more complex process than Part (c). Figures 8 shows an outline which involves recognising the discount must be calculated in five stages and performing calculations represented by the box for each stage. Figure 9 shows in detail the steps involved in one such calculation of the cost for a time interval, which is very similar to Figure 7, although the time interval must be calculated from the graph, rather than used as given from the "10 minutes" cue for Part (c).

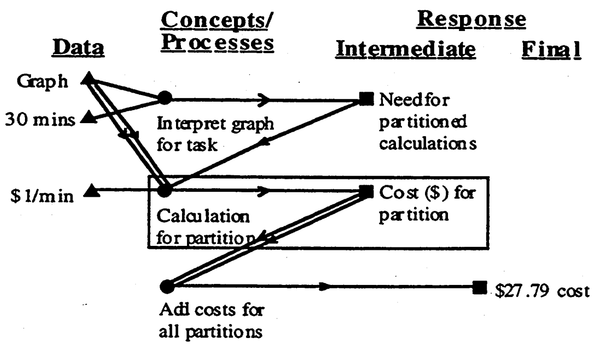


Figure 8. Overview of $27.79 response to Part (d)

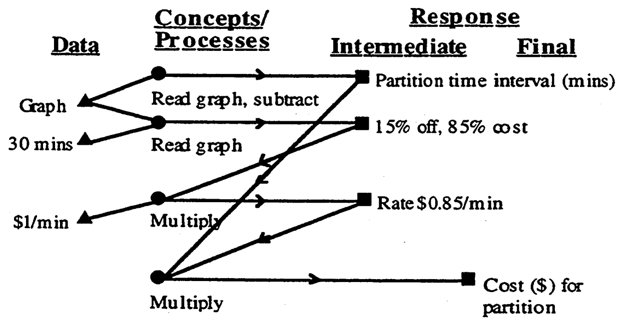


Figure 9. Detail in calculating a partial response to Part (d)

Some of the responses between $20 and $30 were categorised as using a correct approach but involving a calculation error. For example, a number of students responded $23.12. This value is $4.67 less than $27.79 and would result if a systematic error were made in calculating the time interval by subtraction (see Figures 1 and 9), yielding all time intervals one minute less than the correct value. Table 3 shows the percentages of responses in each of the categories above by grade. Again the majority of Grade 6 students responded "$30," apparently seeing the graph as irrelevant to the task. This approach decreased in frequency for older grades but very few were totally successful even at Grade 11.

Table 3
Percentage of Responses to Part (d) by Response Category and by Grade

Response	Grade 6		Grade 8	Grade 9		Grade 11
Category	1993	1995	1995	1993	1995	1995
$27.79	0	0	2	3	2	2
($27.79)	2	1	1	9	4	12
$25.50	4	1	5	8	9	10
15%/$4.50/...	<1	1	3	4	5	7
$30	62	69	46	25	29	21
Other	15	15	15	26	18	21
NR	16	13	28	26	34	27
N	*286*	*216*	*295*	*329*	*279*	*179*

In analysing responses given by the same students longitudinally 1993 to 1995, many students either did not attempt the questions in both 1993 and 1995, or were of indeterminate level (NR/Other), and so were eliminated from longitudinal analysis. Of 114 students who were assigned a level based on the "highest" level for Parts (c) and (d) for both 1993 and 1995, 35 (31%) of students improved, 8 (7%) regressed, and 71 (62%) stayed at the same level. This latter figure was mainly due to 49 Grade 6 students who gave responses consistently ignoring the graph, that is $10 for Part (c) and $30 for Part (d). This suggests that when minor fluctuations of functional performance are taken into account, very few students actually improved their performance over the two year interval.

Discussion

Several important points arise from the responses of students to this item. One dominant finding is the extent to which responses to calculate cost ignored the graph altogether, suggesting that many students could not interpret the graph in context as it applied to Parts (c) and (d). Asking for "cost" may trigger a calculation mode, which excludes consideration of the graph. These students have not reached the second level of statistical literacy in this context. Experiences with handling raw meaningful data may be useful to these students in establishing links between the data itself and representations of the data in different graphical forms.

Another finding is the lack of use of the numeracy skills which are required for completing the tasks in Parts (c) and (d). Many students who did know how to read the graph could only state individual facts based on extracting one aspect or data value from the graph. To go further, it is necessary to understand rate in context, to understand discount, to calculate unusual interval lengths where subtraction of 'endpoints' is invalid, and to have on overview of a multistep procedure. Putting all of these together was very difficult for these students.

While it may be encouraging that about one third of students noticed aspects of the representation which may be unusual, although not technically incorrect, it is somewhat worrying from a statistical literacy point of view that so few noticed the uneven scaling of both axes. There was also a tendency for some students to look for significant patterns in the graphs (cf. Watson & Pereira-Mendoza, 1996) as a way of finding something "unusual". This is not an appropriate methodology in this context.

The results may disturb educators but please advertisers. If between one third and one half of people see nothing wrong with this advertisement, or believe that it "costs less" the longer you talk, it may be that the advertisers have achieved their objective. Students need to be challenged in the classroom using non-standard graphs, even those with errors, to question why the author has represented a message in certain ways and to be on the lookout for misleading representations.

Acknowledgment

This research was funded by the Australian Research Council, Grant No. A79231392. The newspaper advertisement for Telecom which formed the basis for this item appeared in *The Age,* (22 July, 1993, p. 11) and *The Mercury* (22 July, 1993, p. 17).

References

Australian Education Council. (1991). *A national statement on mathematics for Australian schools*. Carlton, Vic. : Author.

Australian Education Council. (1994). *Mathematics - A curriculum profile for Australian schools*. Carlton, Vic.: Curriculum Corporation.

Curcio, F. R. (1987). Comprehension of mathematical relationships expressed in graphs. *Journal for Research in Mathematics Education, 18*, 382-393.

Curcio, F. R. (1989). *Developing graph comprehension*. Reston, VA: National Council of Teachers of Mathematics.

Ministry of Education. (1992). *Statistics in the New Zealand Curriculum*. Wellington, NZ: Author.

Moritz, J. B., & Watson, J. M. (1997). *Pictograph representation: Telling the story*. Paper prepared for presentation at the 16th Biennial Conference of the Australian Association of Mathematics Teachers, Melbourne: July, 1997.

Pereira-Mendoza, L., & Mellor, J. (1991). Students' concepts of bar graphs - Some preliminary findings. In D. Vere-Jones (Ed.), *Proceedings of the Third International Conference on Teaching Statistics. Volume 1: School and general issues* (pp. 150-157). Voorburg: International Statistical Institute.

Pereira-Mendoza, L., Watson, J. M., & Moritz, J. B. (1995). What's in a graph? In A. Richards, G. Gillman, K. Milton, & J. Oliver (Eds.), *Flair: Forging links and integrating resources* (pp. 301-307). Adelaide, SA: Australian Association of Mathematics Teachers.

Watson, J. M. (1994a). A diagrammatic representation for studying problem solving behavior. *Journal of Mathematical Behavior, 13*, 305-332.

Watson, J. M. (1994b). Instruments to assess statistical concepts in the school curriculum. In National Organizing Committee (Ed.), *Proceedings of the Fourth International Conference on Teaching Statistics. Volume 1* (pp. 73-80). Rabat, Morocco: National Institute of Statistics and Applied Economics.

Watson, J. M. (in press). Assessing statistical literacy using the media. In I. Gal & J. Garfield (Eds.), *The Assessment Challenge in Statistics Education*. Voorburg, The Netherlands: International Statistical Institute and IOS Press.

Watson, J. M., Collis, K. F., & Moritz, J. B. (1994). *Authentic assessment in statistics using the media*. Report prepared for the National Center for Research in Mathematical Sciences Education - Models of Authentic Assessment Working Group (University of Wisconsin). Hobart: Uni. of Tasmania, School of Education.

Watson, J. M., & Moritz, J. B. (1996a). *The art of bar charts: Steps in interpreting and predicting*. Poster presented at the International Congress on Mathematical Education, Seville, Spain: July, 1996.

Watson, J. M., & Moritz, J. B. (1996b). *Student analysis of variables in a media context*. A paper presented in Working Group 14 at the Eighth International Congress on Mathematical Education, Seville, Spain: July, 1996.

Watson, J. M., & Pereira-Mendoza, L. (1996). Reading and predicting from bar graphs. *Australian Journal of Language and Literacy, 19*, 244-258.