Using the Media as a Means to Develop Students' Statistical Concepts

Marian Kemp
Head of the Student Learning Centre, Murdoch University
m.kemp@murdoch.edu.au

Abstract

In this era of increasingly fast communication people are being exposed to quantitative information, from national and international sources, through a range of media including newspapers, magazines, television, radio, pod-casts, *YouTube* and other areas of the Internet. Contexts include health statistics, environmental issues, traffic statistics, wars, gun laws and so on. It is becoming more and more important that citizens are able to critically read and interpret this information, and to do so requires an understanding of statistical concepts. Research has shown that students are motivated and engaged in learning through the use of authentic, real life tasks. The media provides current information, which can be used to help develop both students' awareness of how social issues are constructed as well as vital statistical concepts. This paper proposes that secondary school students' application of a model for statistical analysis to material taken from media sources, enhances their understanding of statistical concepts. This model, called the *Five Step Framework*, is described and exemplified for the particular context of opinion polling.

Introduction

Electronic and digital communication across the globe has brought with it more opportunities for transmission of information and the need for people to be statistically literate; to be able to make sense of the information presented to them in a variety of ways and formats. This paper will concentrate on the development of statistical concepts through the analysis of data from media sources. As Ben-Zvi and Garfield (2004) point out:

> Quantitative information is everywhere, and statistics are increasingly presented as a way to add credibility to advertisements, arguments or advice. Being able to properly evaluate evidence (data) and claims based on data is an important skill that all students should learn as part of their educational programs. The study of statistics provides tools that informed citizens need in order to react intelligently to quantitative information in the world around them. Yet many research studies indicate that adults in mainstream society cannot think statistically about important issues that affect their lives (p.3).

Given the indications that citizens are often unable to engage with the quantitative information around them, it is strongly suggested here that students be exposed to the interpretation and analysis of information from the real world, presented via the media while still at school. Interpreting material in the media provides a number of benefits to the students. Firstly, it provides authentic tasks for which students have to interpret the statistical language, deal with incomplete information, develop problem-solving strategies, do some research and make decisions for themselves. These authentic tasks, as advocated by Zevenbergen (1997), can engage and motivate students in a variety of contexts.

The literature indicates some agreement about the statistical concepts and skills that students should develop during their years at school. These include (i) developing appropriate research questions, (ii) collecting, organising and displaying data, (iii) selecting appropriate analysing tools, (iv) understanding probability in the context of uncertainty, (v) making inferences and predictions (e.g., Watson, 2006; Ben-Zvi & Garfield, 2004). There are pedagogically sound activities that can be undertaken in the classroom (e.g., Curriculum Council, 2009; Watson, 2004) to enable students to develop statistical concepts and skills. The author stresses here the importance of the integration of examples from the media in the process of students developing their ability to critically evaluate data in the media used in 'advertising, arguments or advice'. Given the space constraints, the paper will focus on the interpretation of data from opinion polls that are really prevalent in the media and usually reported in tabular form.

Description of the model and its adaptation to opinion polling

This paper proposes the use of a specific, tried and tested, model that appreciates people's need to interpret the statistical concepts and language that they meet in their everyday lives. It suggests that exposure to data from a range of media outlets can help students develop their statistical

concepts and at the same time helps students to learn how to critically evaluate the information they are analysing, which is recognized as especially important (Steen, 1997).

This model is a *Five Step Framework,* designed to interpret tables and graphs (Kemp, 2005). The early development of the model, shown in Table 1 below, was reported in Kemp (2003). The framework was designed and refined to cater for the needs of university students in their study where they need to interpret data provided in academic journals, scientific reports, textbooks, lectures, media articles and other materials. The framework has been used successfully in workshops for science students in a first year statistics unit (Kemp & Bradley, 2006) and with pre-service primary teachers (Kemp, 2005). Pre- and post-tests before and after workshops for both sets of students showed that the strategies built through the *Five Step Framework* significantly increased students' ability to interpret tables of data. The framework has been successfully used with other students in Foundation Units and Alternative Entry Programs at Murdoch University, but the effectiveness has not been formally evaluated and reported.

Table 1. Five Step Framework (Kemp, 2005).

Step 1: Getting started Look at the title, axes, headings, legend, footnotes and source to find out the context and expected reliability of the data.
Step 2: WHAT do the numbers mean? Make sure you know what all the numbers (percentages, '000s etc) represent. Look for the largest and smallest values in one or more categories or years to get an idea of the range of the data.
Step 3: HOW do they change or differ? Look at the differences in the values of the data in a single data set, a row, column or part of a graph. Repeat this for other data sets. This may involve changes over time, or comparisons within categories, such as male and female, at any given time.
Step 4: WHERE are the differences? What are the relationships in the table or graph? Use your findings from Step 3 to help you make comparisons between columns or rows in a table or parts of a graph to look for similarities and differences.
Step 5: WHY do they change? Look for possible reasons for the relationships in the data you have found by considering societal, environmental and economic factors. Think about sudden or unexpected changes in terms of state, national and international policies or major events.

This paper focuses, however, on the education of secondary students and the integration of the use of media materials into their curriculum. It is really important that students appreciate that, especially in opinion polling, a critical element of reliability of the data concerns potential bias, for which there are at least three different possible sources.

The first way in which an opinion poll can be biased is directly related to the question(s) asked. These can be leading questions that can persuade the interviewee to give an answer in accord with the interviewer's point of view. For example, the question "Do you believe that our boys and girls should be withdrawn from Iraq?" poses a different question from "Do you think that Australia has a moral obligation to protect the oppressed in Iraq?" Even though both questions relate to whether people think that Australian troops should stay in Iraq, or be withdrawn, they are likely to elicit quite different responses.

Secondly, for the data from any poll to be valid, the sample needs to be a random representative sample. In this respect there is the need for samples to be taken at random, (where random sampling is a method of selecting a sample in which all possible subjects (or scores) in the population have an equal chance of being selected), and all possible samples have an equal chance of being selected. This can be done through using random number generators of some kind or putting names in a proverbial hat. By this mechanism of random sampling we hope that the sample will be unbiased and that the sample will include the same characteristics as the population. However, as Baldi and Moore (2009) point out "[r]andom samples eliminate bias from the act of choosing a sample, but they can still be wrong because of the variability that results when we choose at random (p. 220)". Thus, it can happen that when a population has a very wide variability random samples can give quite different results.

Thirdly, the size of the sample also contributes to the credibility of the results. Essentially, the sample size must be determined to give a sufficiently small standard error. Noting Baldi and Moore's (2009) comments above, account should be taken of the variability of the different groups within the population to increase the chance that the range of opinions is captured and the sample is

not biased by under- or over-representation of particular subgroups.

Therefore, it follows that in analysing opinion poll data it is essential that students explicitly address the three questions below:

1. What was the question asked?

(What did the researcher want to know? Was the question biased?)

2. What was the method of sampling and data collection?

(Was it a random sample? How was the sample collected and collated?)

3. What was the sample size?

(How many respondents are in the sample? How would this affect the variability of the results?)

These questions are appropriately located in Step 1 of the *Five Step Framework*, expanding the question about reliability referred to in Table 1. Step 1 in Table 3 below includes suggested sequences of questions to help students explore the data. Different levels of students' age and level of background should dictate the level of complexity of the data they are analysing. For simpler opinion polls and less sophisticated students it is probably more appropriate to concentrate on Steps 1 to 3. It is expected that for older students all five steps can be accomplished.

To illustrate the use of this model, consider an opinion poll conducted recently in Western Australia concerning daylight saving, This has been a topic of continued debate in the West Australian community over the last three years of a trial of daylight saving. The debate has included a wide range of perspectives including those of farmers, businessmen, surfers and families with young children. Table 2 represents a table of data published in *The West Australian* daily newspaper on 11 April 2009 in an article headed 'On a Knife Edge' (Phillips, 2009).

Table 2. Opinion poll results concerning daylight saving

Westpoll Do you support daylight saving?

	March '07	March '09	April '09
Support	34	42	47
Oppose	62	57	51
Don't know	4	1	2

Westpoll conducted April 6-8 through phone interviews with 400 voters across WA by Patterson Market Research

Table 3, which follows, indicates that Step 1 has been extended to take account of the questions described above. The questions in italics are those that have been routinely asked for any table of data.

As can be seen in Table 3 the five steps are amplified with typical questions that would be asked in the interpretation of any table of data. Once students have developed the strategies needed to interpret a table of data, after using the framework a few times, they can generate their own questions to explore the data. Emphasis has been given to Step 1 to indicate the importance of addressing the three questions described above. For some students the steps that are important to them in this context will be Step 1 to Step 3, while more sophisticated students would be encouraged to complete all five steps. It is vital to establish the links between the concepts and skills that students are learning in the classroom and data they explore in the media. Naturally, students can go to the website of Patterson Market Research (2009), or discuss with other students and their teacher to understand the significance of the potential bias in the sampling.

Applying the *Five Step Framework* to a more complex table

The context of an opinion poll on identity cards provides a much more complex table for analysis. Currently in Australia it is not a requirement for people to have an identity card and a survey was conducted in 2006 to investigate public opinion on this issue. The poll was conducted by *Newspoll* and the results published in the 27-28 January edition of *The Australian* daily newspaper. The table above is a reproduction of the data presented in the newspaper.

In this case the data is far more complex and involves more sophisticated comparisons to extract the meaning. The *Five Step Framework* gives students the structure to develop the strategies for interpreting this table of data. Naturally the interrogation of this table would be done by students who cannot only complete Steps 1 to 3 but also to Step 4 to examine the complex relationships in

this table.

Table 3. Application of the Five Step Framework to Table 2 on Daylight Saving

Step 1: Getting started
Q: *From the title, what is the general topic being examined?*
Q: How are the variables being compared?
From the labels on the left column, how are the groups being compared?
From the labels on the top row, how are the groups being compared?
Q: What question was asked?
Do you think the question was biased?
Q: According to the table, over what timeframe was the poll undertaken?
Which data would have been collected at that time?
Are the collection dates evenly spread?
When were the data for March 2007 and March 2009 collected?
Q: *Is there any evidence to suggest that the data is reliable?*
How was the sample selected?
Who selected the sample?
Was it a random selection?
Q: How was the data collected?
How was the phone poll organised?
Were only landline phones included?
How many people in each household were interviewed and how were they chosen within the household?
How did the interviewer check that the interviewee was on the electoral roll?
Q: What was the size of the sample?
Is this is a reasonable number to gain a sampe of appropriate size?
Q: What else would you like to know about the sampling?
Spend some time researching this data collection Market Research Agency.
Step 2: WHAT do the numbers mean?
Q: What is the meaning of the 47 in the first column?
Q: Which year has the highest percentage of people in favour?
Q: Which category has the lowest percentage of people in favour?
Step 3: How do they change?
Q: How does the percentage of people in support change over the three collection times?
Q: How does the percentage of people who oppose change over the three collection times?
Q: How does the percentage of people who don't know change over the three collection times?
Step 4: WHERE are the differences?
Q: Compare the differences for 'in favour' and 'against' for each time frame, look at the relationship between the percentages.
Q: Compare the don't know responses with the support percentages.
Step 5: WHY do they change?
Q: Did any of the values in the table surprise you?
Q: Suggest possible reasons for the differences in the percentages for the responses considering that two of the dates are 3 years apart and the third is only one month later than the second.

Table 4. National identity card

Question: Are you in favour or against the introduction of a national identity card in Australia? If in favour - is that srongly in favour or somewhat in favour? If against-is that strongly against or somewhat against?

January 27-29, 2006

		Sex			**Age**		**Political**	**Support**
Percentage Support	Total	Male	Female	18-34	35-49	50+	Coalition	Labour
Strongly in favour	27	30	23	13	25	37	31	25
Partly in favour	26	24	29	28	26	26	29	27
Total in favour	**53**	**54**	**52**	**41**	**51**	**63**	**60**	**52**
Strongly against	12	11	14	16	12	10	14	11
Partly against	19	23	15	20	23	16	13	23
Total against	**31**	**34**	**29**	**36**	**35**	**26**	**27**	**34**
Uncommitted	16	12	19	23	14	11	13	14

This survey was conducted on January 27-29, 2006 on the telephone by trained interviewers in all states of Australia and in both city and country areas among 1200 people aged 18 years and over. Telephone numbers and the person within the household were selected at random. The data has been weighted to reflect the population distribution. The maximum margin of sampling error in the total sample is plus or minus 3 percentage points. More poll information is available at www.newspoll.com.au.

.

There is a wide range of useful comparisons that can be made across the categories of age, gender and political persuasion and the levels of agreement or disagreement. In order to cope adequately, students need a strong grasp of proportions and the author found that special tutorials were needed for some students to develop those concepts and skills adequately to work with more complex tables. In relevant situations, particularly those involving social issues, the ultimate aim is to enable students to reach and successfully complete thoughtful answers to the kinds of questions posed in Step 5, which involve locating the data in terms of national and international issues.

Conclusion

Through the incorporation of material from the media in the classroom, students' understanding of statistical concepts can be enhanced. It has been observed from experience that without models of the kinds described in this paper, many students will not be able to interpret media reports. Students need support to develop the strategies necessary to do so. This paper has shown how a model can be adapted successfully for this purpose, both for unsophisticated and more complicated tables. It has focused on the particular needs concerned with analysing the data from opinion polls. Emphasis has been placed on consideration of the sample selection and collection. This seems to be best placed in Step 1 for all students. Tables with a simple structure can be tackled by younger students whereas more complex tables need an understanding of proportion and other means of comparison, more suited to more sophisticated students. Thus, not all students will reach Steps 4 and 5 but the aim over time is to enable students to become adept with thinking about relationships between the variables, and ultimately about why the data are as they are and how they relate to society in a range of ways.

References

Baldi, B., & Moore, D. S. (2009). *The Practice of statistics in the life sciences*. New York: W.H. Freeman and Company.

Ben-Zvi, D., & Garfield, J. (2004). Statistical Literacy, Reasoning, and Thinking: In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking.* (pp. 3-16). The Netherlands: Kluwer Academic Publishers.

Curriculum Council. (2009). *Chance and Data Resources Stages 2 and 3*. Retrieved April 26, 2009, from http://www.curriculum.wa.edu.au/internet/Senior_Secondary/Courses/Mathematics.

Kemp, M. (2003). Critical numeracy: Helping people to decide. In A.Rogerson (Ed.). *Mathematics into the 21st Century Project Conference. Proceedings of The Decidable and the Undecidable in Mathematics Education*. Brno: Czech Republic. (pp.144-148) Retrieved May 31, 2009, from http://math.unipa.it/~grim/21_project/21_brno03_kemp.pdf

Kemp, M. (2005). *Developing critical numeracy at the tertiary level*. Digital thesis, Murdoch University Library. Retrieved May 5, 2009, from http://wwwlib.murdoch.edu.au/adt/browse/view/adt-MU20060831.171947

Kemp, M. (2009). Numbers in the Media. In C. Hurst, M. Kemp, B. Kissane, L. Sparrow & T. Spencer (Eds.), *Mathematics: It's Mine: Proceedings of the 22nd Biennial Conference of the Australian Association of Mathematics Teachers*. (pp. 114-124) Adelaide: Australian Association of Mathematics Teachers, Inc.

Kemp, M, & Bradley, B. (2006). Developing statistical literacy across the university curriculum. In D. Quinney, D. (Ed.), *International Conference on Teaching Mathematics*, Istanbul, Turkey, July 2006. (Electronic proceedings, paper No 221. Retrieved 31 May, 2009, from http://www.tmd.org.tr/ictm3/readme.html

Patterson Market Research. (2009). *Daylight saving poll*. Retrieved May 3, 2009, from http://www.marketresearch.com.au/our_services__westpoll.5.html

Phillips, J. (2009). On a knife edge. *The West Australian*, (11 April 2009, p.1).

Steen, L. A. (Ed.). (1997). *Why numbers count: Quantitative literacy for tomorrow's America*. New York: The College Board.

Watson, J. (2004). Developing reasoning about samples. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking*. (pp. 277-294). The Netherlands: Kluwer Academic Publishers.

Watson, J. (2006). Sampling –A good start. *Statistical Literacy at School: Growth and Goals* (pp. 27-54). New Jersey: Laurence Erlbaum Associates Inc.

Zevenbergen, R. (1997). Situated numeracy: A case study of pool builders. In N. Scott & H. Hollingsworth (Eds.), *Proceedings of the 16th Biennial Conference of the Australian Association of Mathematics Teachers, Mathematics creating the future* (pp. 89-94). Melbourne: AAMT.