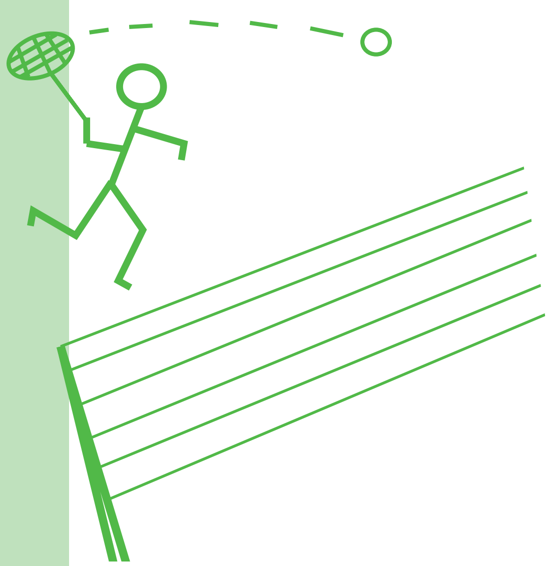
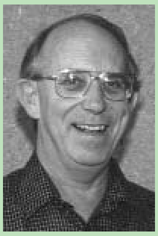


The role of context in students' analysis of data



Steven Nisbet, Cynthia Langrall and Edward Mooney explore the impact of students' knowledge of real-life events has on the way they analyse data.

What role does students' knowledge of real-world contexts play in their analysis and interpretation of data? This study found that primary-aged students used context knowledge in three broad ways: to rationalise the data or their interpretations, in taking a critical stance toward the data, and in ways that were not necessarily productive or pertinent in addressing the task at hand.

When teaching students how to analyse data, it is important to realise that data are numbers in context and that data engage our knowledge of the context so that we can understand and interpret rather than simply carry out arithmetical operations (Moore, 1990). By context we mean the real-world phenomena, settings, or conditions from which data are drawn, and in this study we investigated how students draw on their context knowledge when analyzing data.

Our study was informed by a number of theoretical considerations. First, Dapueto and Parenti's (1999) theoretical model describes the relationship between context and the formation of mathematical knowledge, and from that theory we identified three relevant factors for our study: (i) the students' field of experience, or familiarity with the context of data being analyzed, (ii) whether the data analysis or interpretation necessitates the use of certain statistical knowledge, and (iii) the meaningful role statistics plays in understanding or interpreting the data. Secondly, the notion of shifting or the interplay between data and context has been described by a number of researchers (e.g., Pfannkuch & Wild, 2004; Shaughnessy, Garfield, & Greer, 1996; Watson & Callingham, 2003). Thirdly, Gal's (2004) model of statistical literacy describes the types of knowledge (literacy skills, statistical knowledge, mathematical knowledge, context knowledge, and critical questions) and dispositions (beliefs, attitudes, critical stance) that enable a person to comprehend, interpret, critically evaluate, and react to statistical messages.

The students

We worked with six students (3 males and 3 females, aged between 11 and 12) from a large primary school in a middle-class Brisbane suburb. The classroom teacher selected the students in response to our request for two groups of three students such that one student in each group had great interest in and expert knowledge of some particular topic or pursuit, and hence could be described as having extensive contextual knowledge. The other four students were to have no great interest in the first two students' topics. The groups were comprised as follows: Group A – Larry (tennis fanatic), Tracey, and Don; and Group B – Beth (a female pop-singer fanatic), Mandy, and Joel. All names are pseudonyms.

The tasks

After the groups were selected and the areas of expertise identified (tennis champions and pop stars), we searched for sets of data which would form the basis of the tasks. Finding relevant data sets was not difficult, given the power of searching on the web. Armed with our newly-found data sets, we constructed tasks that required students to make comparisons between data sets. Two tasks corresponded with the two "expert" students' areas of interest (tennis champions and pop singers) to help us examine the role that context knowledge played in analysing the data. A third task was developed on a topic that was thought to be of no special interest to any of the students, namely, champion discus throwers. Again, finding relevant data on the internet was relatively easy. The three tasks are now described. Note that each task was framed in terms of a problem, to engage students' higher-order thinking as they interact with the data.

The complete data sets can be found on the following website: www.math.ilstu.edu/langrall.

Session 1 task: Tennis Players

Does the USA produce better men's tennis players than Australia?

Two tables of data were presented; (a) a listing of the winners and runners-up of the men's Wimbledon championship from 1955 to 2004, including the players' nationality and seedings, the final score, and the match time; and (b) a listing of the top 20 male tennis players with their current ranking, nationality, prize money for 2005 to date, and career prize money. Parts of the data sets are given in figures 1 & 2.

Figure 1: Wimbledon champions

Year	Champion	Seed	Runner-Up	Seed	Score	Mins
2004	R. Federer (SUI)	1	A. Roddick (USA)	2	4-6, 7-5, 7-6 (7-3), 6-4	150
2003	R. Federer (SUI)	4	M. Philippoussis (AUS)	U	7-6 (7-5), 6-2, 7-6 (7-3)	116
2002	L. Hewitt (AUS)	1	D. Nalbandian (ARG)	28	6-1, 6-3, 6-2	117
2001	G. Ivanisevic (CRO)	U	P. M. Rafter (AUS)	3	6-3, 3-6, 6-3, 2-6, 9-7	181
2000	P. Sampras (USA)	1	P. M. Rafter (AUS)	12	6-7 (10-12), 7-6 (7-5), 6-4, 6-2	182
1999	P. Sampras (USA)	1	A. K. Agassi (USA)	4	6-3, 6-4, 7-5	115
1998	P. Sampras (USA)	1	G. Ivanisevic (CRO)	14	6-7 (2-7), 7-6, 6-4, 3-6, 6-2	172
1997	P. Sampras (USA)	1	C. A. Pioline (FRA)	U	6-4, 6-2, 6-4	94

Figure 2: ATP Prize Money

Player	Rank	Nationality	Prize money US$ 2005 (to 18 April)	Career prize money (US$)
Agassi, Andre	7	USA	372,062	29,738,741
Ancic, Mario	19	CRO	228,420	1,213,950
Canas, Guillermo	18	ARG	230,787	3,546,047
Coria, Guillermo	8	ARG	339,892	4,682,517
Davydenko, Nikolay	10	RUS	327,933	1,884,270
Federer, Roger	1	SUI	1,772,078	15,867,633
Ferrer, David	9	ESP	336,816	1,271,840
Gaudio, Gaston	13	ARG	265,945	4,157,833
Gonzalez, Fernando	16	CHI	240,275	2,692,620
Henman, Tim	20	GBR	228,100	10,845,240
Hewitt, Lleyton	4	AUS	756,955	15,259,975

Session 2 task: Pop Singers

In a recent pop culture survey, teenagers identified Britney Spears and Delta Goodrem as two of the top female performers. Based on the following data collected from the Top 40 Charts, which of these two singers is more popular?

A table of data (Figure 3) was presented for each singer listing seven "Top 40" hits between 2002 and 2005; song title, total number of weeks on the charts, dates and positions for five highest rankings for each song.

Does this extra information help you interpret the data?

Tables were presented with additional data (Figure 4) on total number of different charts and specific charts for top five rankings.

Figure 3: Delta Goodrem Top 40 Hits (2002–2005)

Song Title	Total Weeks on Charts/ Number of Different Charts	Top Appearances	Date	Position
Born to Try	130 weeks / 10 charts	Australia Top 40 New Zealand Top 40 Australia Top 40 Australia Top 40 Australia Top 40	30-11-2002 18-05-2003 23-11-2002 07-12-2002 14-12-2002	1 1 2 2 2
Lost without You	147 weeks / 12 charts	Australia Top 40 Australia Top 40 Australia Top 40 Australia Top 40 Australia Top 40	08-03-2003 22-03-2003 15-03-2003 19-03-2003 03-05-2003	1 1 2 2 2
Innocent Eyes	76 weeks / 8 charts	Australia Top 40 Australia Top 40 Australia Top 40 Australia Top 40 UK Top 20	28-06-2003 05-07-2003 14-05-2003 21-06-2003 06-07-2003	1 1 2 2 2

Figure 4: Britney Spears Top 40 Hits (2002–005)

Song Title	Total Weeks on Charts/ Number of Different Charts	Top Appearances	Date	Position
I'm not a Girl, Not Yet a Woman	158 weeks / 19 charts	UK Singles Top 40 Europe Official Top 40 Austria Top 40 Europe Official Top 40 Ireland Top 20	07-04-2002 11-04-2002 12-04-2002 18-04-2002 27-04-2002	2 2 3 3 3
I Love Rock and Roll	103 weeks / 16 charts	Portugal Top 40 Portugal Top 40 German Top 40 Portugal Top 40 Ireland Top 20	24-07-2002 31-07-2002 08-06-2002 17-07-2002 09-11-2002	4 5 7 8 8
Boys	99 weeks / 21 charts	World Jazz Top 20 Singles World Jazz Top 20 Singles UK Singles Top 40 Belgium Top 20 Belgium Top 20	08-08-2002 15-08-2002 04-08-2002 07-09-2002 28-09-2002	1 6 7 7 9

Session 3 task: Discus Distance

These tables (Figure 5) show the winning discus throws for men and women in the Summer Olympics from 1928 to 2000. Based on these data, who has shown better performance over the years – men or women?'

Two tables were presented, one for men and one for women. Each table included 3 columns of information: year, winner's name, and distance in metres.

Examine this graph. What do these data tell you?

A double bar graph representing the same data as the table in Figure 5 was presented

Figure 5: Winning Discus Throws in Summer Olympics

Men's Discus Winners		
Year	Name	Distance (metres)
1928	Bud Houser, USA	47
1932	John Anderson, USA	49
1936	Ken Carpenter, USA	51
1948	Adolfo Consolini, ITA	53
1952	Sim Iness, USA	55
1956	Al Oeter, USA	56
1960	Al Oeter, USA	59
1964	Al Oeter. USA	61
1968	Al Oeter, USA	65
1972	Ludvik Danek, CZE	64
1976	Mac Wilkins, USA	67
1980	Victor Rashchupkin, USSR	67
1984	Rolf Danneberg, W. GER	67
1988	Jurgen Schult, E. GER	69
1992	Romas Ubartas, LIT	65
1996	Lars Riedel, GER	70
2000	Virgilijus Alekna, LIT	69

Women's Discus Winners		
Year	Name	Distance (metres)
1928	Halina Konopacka, POL	40
1932	Lillian Copeland, USA	41
1936	Gisela Mauermayer, GER	48
1948	Micheline Ostermeyer, FRA	42
1952	Nina Romaschkova, USSR	52
1956	Olga Fikotova, CZE	54
1960	Nina Ponomaryeva, USSR	55
1964	Tamara Press, USSR	57
1968	Lia Manoliu, ROM	58
1972	Faina Melnik, USSR	67
1976	Evelin Schlaak, E. GER	69
1980	Evelin Schlaak Jahl, E. GER	70
1984	Ria Stalman, NED	65
1988	Martina Hellmann, E. GER	72
1992	Maritza Marten, GER	70
1996	Ilke Wyludda, GER	70
2000	Ellina Zvereva, BLR	68

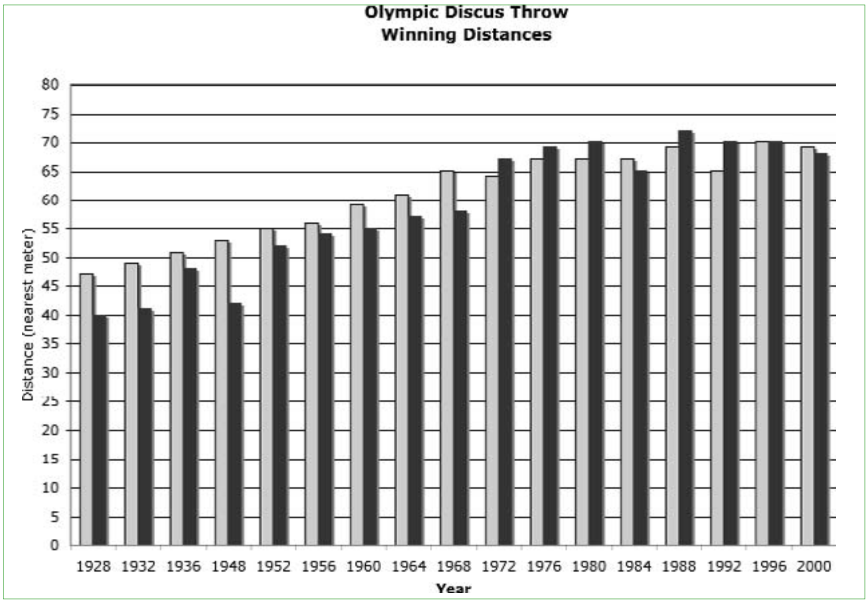


Figure 6: Graph of data for Winning Discus Throws

Each session lasted about 45 minutes and started with a brief whole-group discussion during which the students were introduced to the task. The students then worked in their assigned groups, and one of the two researchers monitored their discussion and encouraged students to contribute their thoughts and suggestions for analyzing the data. We refrained from leading the discussion in any particular direction although we did ask probing questions to better understand students' thinking as they engaged in the task and when necessary to refocus their attention on the task at hand. Each session closed with another whole-group discussion during which students from each group shared their interpretations of the data.

Both groups were video- and audio-taped, and the students were encouraged to record their answers and opinions on the sheets of paper provided. We used qualitative methods of analysis to identify and characterize students' use of statistical knowledge and context knowledge and the interplay between them. We were particularly interested in how students shifted between context knowledge and statistical knowledge, as in this episode from Group A's discussion of Wimbledon Champions whilst tallying the number of winners from Australia and USA. As they count the totals, Larry interjects: "But the only problem is most of the [Australian] winners and runners up, all of them were back before 1980 so they don't play tennis anymore. Connors and McEnroe still play a little bit; they play in the oldie's tour. . . . So let's just look at the 1990s. For the champions, you've got one Australian…". In the first part of this excerpt, Larry made a comment based on his reading of the data noting that most of the Australians who won Wimbledon championships did so prior to 1980. He transitioned seamlessly into a discussion about the players who are no longer active on the tennis circuit and those who continue to play "in the oldie's tour." We consider this shift to reflect the use of context knowledge because the data set does not indicate whether the players are still active. Based on this knowledge, Larry suggested they partition the data set and examine information only for the 1990s and later, shifting back to a statistical focus.

The experts' use of context knowledge

For both Task 1 (tennis champions) and Task 2 (pop stars), the group with the expert for the task at hand engaged in a greater percentage of discussion episodes related to context than the group without an expert. When the group included an expert for the task at hand, about 50% of the episodes reflected the use of context knowledge. By contrast, in the group without an expert, only 25% of the episodes related to the context. Our findings also show that for Tasks 1 and 2, the respective expert initiated the majority of context-related episodes. For example, with the Tennis Task, 11 of the 13 context-related episodes (85%) were attributed to Larry, the tennis expert. With the Pop-Star Task, 6 of the 8 context-related episodes (75%) were attributed to Beth, the expert on pop singers.

For Task 3 (Discus throwing), for which neither group included an expert, the percentages of episodes reflecting the use of context knowledge were much less (31% for Group A and 35% for group B). Further, the distribution of students who initiated context-related episodes was generally balanced within each group.

How the students used context knowledge

The students used context knowledge three ways. The most common use was to rationalize the data or support their own interpretation of the data. This occurred in more than half of the context-related episodes across both groups. For example on Task 3, the students in Group B decided to focus on discus throwers with winning distances greater than 50 metres. Having identified that only two men had distances less than 50 metres, Mandy and Beth offered the following explanation:

M: Maybe see how in the earlier years, maybe the discus, the disc thingy was heavier because they didn't have the technology that made it lighter. It's big and in the earlier years it was pretty heavy, but now it's a big one and it's lighter than the big one in the earlier years.

B: And maybe a different type of ground and they didn't have the right shoes or something.

The second and third uses of context each occurred in about 23% of the episodes. One of these uses of context was characterized as being nonproductive to the analysis of data or not pertinent to the task at hand. For example, on Task 2 Beth initiated episodes that focused on the immaturity of Britney Spears and the fact that Delta Goodrem now has a boyfriend. The other use of context, however, reflected what we consider to be a desirable approach to data analysis and interpretation. It involves taking a critical stance that includes (a) questioning how the data were collected or what they represent; (b) recognizing the need for more or different data; and (c) bringing new information or insight to the task at hand. For example, on Task 1, Larry initiated two episodes in which he commented on the need for more information about the results of all of the grand slam tennis tournaments, not just Wimbledon.

We conclude that knowledge of context is an important factor contributing to students' engagement in statistical tasks. At times, knowledge of the context can assist in data analysis and positively influence the discussion within a group, but, at other times, context knowledge can be distracting and of little assistance in completing the task at hand. As might be expected, we found that the context expert tended to dominate context-related discussions. We also found that the context expert was able to establish authority within the group in that his or her expertise was recognized and upheld by the group members.

How students used statistical knowledge

The presence of an expert with respect to context appeared to have no effect on the statistical knowledge brought to bear on the three tasks in this study. Across tasks, episodes reflecting the use of statistical knowledge were attributed to a mix of students, with no one student dominating. Statistical concepts and skills used by the students included the following: reading the data in tables; visually inspecting the data in graphs; making point-wise comparisons of data values; partitioning data into subsets; describing trends in the data; computing averages; making predictions based on the data; and identifying greatest, least, and extreme values in the data.

Conclusions

Although some of the findings of this study may not come as any surprise to experienced teachers, we believed that it was important to investigate these issues and document the results. The presence of a context expert with respect to the data influenced students' work, specifically increasing the percentage of episodes reflecting the use of context knowledge. Even without particular expertise, all three uses of context knowledge were evident in students' deliberations. However, some uses of context were more helpful and relevant than others in terms of completing the task at hand. Interestingly, although context-related discussions were sometimes protracted, we never found more than two consecutive episodes reflecting the use of context knowledge. This was not the case with the use of statistical knowledge. We found that students employed a variety of strategies in analysing the data, resulting in multiple episodes reflecting the use of statistical knowledge. Nevertheless, there was a periodic shifting between the use of context and statistics, supporting our theoretical view of a potential interplay between context knowledge and statistical knowledge. The findings of this study emphasize the importance of providing opportunities for students to engage in investigations that encourage them to integrate contextual and statistical information (Pfannkuch & Wild, 2004; Watson & Callingham, 2003) in order to develop statistical literacy (Gal, 2004).

In terms of pedagogical implications, teachers should be aware of students' context knowledge, and be skillful in forming groups for data analysis and interpretation. They should recognize that context experts may contribute positively to the discussion, but may also bring up issues that can distract the group or dominate and stifle contributions from other group members. Teachers should note that students respond differently to different tasks and different contexts, according to the extent of their interest and expertise.

Teachers should also be conscious of the fact that tasks such as those described in this study provide good opportunity for students to (a) engage with and interpret real-world data, and (b) mathematize problem situations. Teachers therefore should know their students' interest areas and actively seek out related data sets. They should take advantage of the power of searching the internet to find interesting data sets. The tasks described in this study are open-ended problems and give a wide range of students (ability wise) the opportunity to respond at their respective levels. The use of such tasks also will reinforce, for students, a view of mathematics as a relevant, interesting and motivating subject.

Acknowledgement:

This article is based on data presented in a research paper at the 7th International Conference on Teaching Statistics (ICOTS-7) in July 2006 in Salvador, Brazil.

References

Dapueto, C. & Parenti, L. (1999). Contributions and obstacles of contexts in the development of mathematical knowledge. *Educational Studies in Mathematics, 39*, 1–21.

Gal, I. (2002). Adult's statistical literacy: Meaning, components, responsibilities. *International Statistical Review, 70*, 1–52.

Moore, D. (1990). Uncertainty. In L. Steen (Ed.), *On the Shoulders of Giants: New approaches to numeracy* (pp. 95–137). Washington, D.C.: National Academy Press.

Pfannkuch, M., & Wild, C. (2004). Towards an understanding of statistical thinking. In D. Ben-Zvi & J. Garfield (Eds), *The Challenge of Developing Statistical Literacy, Reasoning, and Thinking* (pp. 17–47). Dordrecht: Kluwer.

Shaughnessy, M. J., Garfield, J. & Greer, B. (1996). Data handling. In A. J. Bishop, K. Clements, C. Keitel, J. Kilpatrick & C. Laborde (Eds), *International Handbook of Mathematics Education, Part I* (pp. 205–238). Dordrecht, The Netherlands: Kluwer.

Watson, J. & Callingham, R. (2003). Statistical literacy: A complex hierarchical construct. *Statistics Education Research Journal, 2*, 3–46.

Steven Nisbet
Griffith University
<s.nisbet@griffith.edu.au >
Cynthia Langrall
Illinois State University
<langrall@ilstu.edu>
Edward Mooney
Illinois State University
<mooney@ilstu.edu>

APMC