Exploring Fijian High School Students' Conceptions of Averages

Sashi Sharma, School of Education, University of Waikato, New Zealand

Abstract

This paper focuses on part of a much larger study that explored form five (14 to 16 year-old) students' ideas in statistics. A range of ideas was explored, including the students' ideas about measures of centre and graphical representations. Students' ideas about measures of centre were analysed and categories of responses identified. While students could compute mean and median, they were less competent with tasks that involved constructing meanings for averages. This could be due to an emphasis in the classroom on developing procedural knowledge or to linguistic and contextual problems. Some students used strategies based on prior school and everyday experiences. The paper concludes by suggesting some implications for mathematics education.

Introduction

Despite statistics being a relatively new discipline (Shaughnessy & Pfannkuch (2004), it has gained increased attention in our society in the last 15 years. Many everyday activities often require an understanding of statistics to make intelligent decisions. Decisions concerning business, industry, employment, sports, health, law and opinion polling are made using an understanding of statistical information and technique (Wallman, 1993). However, often the data that people see is tainted (Scheaffer, Watkins, & Landwehr, 1998). So what can be done? In response to this question, there has been a movement in many countries to include statistics at every level in the mathematics curricula. In western countries such as Australia (Australian Education Council, 1991), New Zealand (Ministry of Education, 1992, 2007) and the United Kingdom (Holmes, 1994) these developments are reflected in official documents and in materials produced for teachers. In line with these moves, Fiji has also produced a new mathematics prescription at the primary level that gives more emphasis to statistics (Fijian Ministry of Education, 1994). As statistics education continues to mature as a discipline, statistics educators are paying more attention to developing statistical thinking. This shift in statistics education means emphasis on teaching from statistical techniques, formulas and procedures to developing statistical reasoning and thinking (Bakker, 2004, Gal, 1998).

Many statistics educators (Cai, 1998; Gal, 1995; Konold & Pollatsek, 2002; Mokros & Russell, 1995; Shaughnessy, 2006; Watson & Moritz, 2000) claim that averages play a central role in statistics. For instance, Cai (1998) and Gal (1995) believe that the arithmetic average is not only a core concept in statistics but also an important concept for making informed decisions. Cai adds that statistical analysis and inferences are conducted almost exclusively through the determination of measures of centres, such as the mean. According to Mokros and Russell (1995), one goal of statistics is "to reduce large unmanageable and disordered quantities of information to summary representations" (p. 20). They believe that averages are tools that can be used to make sense of a data set and, in conjunction with the standard deviation, could be used to summarise and compare data sets. Shaughnessy (2006) accentuates that summary representations of centres play a descriptive role or an inferential role in statistics. He concurs with Mokros and Russell (1995) that 'means' and other measures of centres can help summarise information about a data set or an entire sample. Additionally, he explains that if the data has been appropriately collected from a parent population, the sample mean might provide some useful information about the population mean.

Watson and Moritz (2000) reported that for a long time, the term *average* has been synonymous with the arithmetic mean in the school curriculum, and that until the middle of last century the mean was still the measure of centre discussed in mathematics books. They claim that only in the last decade have the three measures of averages (mean, median and mode) been acknowledged in the school curriculum in a definitive way. In New Zealand (Ministry of Education, 2007), the new curriculum document stresses the importance of statistical investigations and statistical literacy in relation to using measures of centre and evaluating choice of measures respectively. Research (Cai, 1998; Konold & Pollatsek, 2002; Mokros & Russell, 1995; Watson & Moritz, 2000) shows that many students find averages difficult to learn and understand in both formal and everyday contexts and that learning and understanding may be influenced by ideas developed in early years. Konold and Pollatsek (2002) express the view that current teaching is not helping students to develop measures of centre. They claim that although many students know how to compute means or medians, they do not know how to apply or interpret these ideas. These views are consistent with the findings of other researchers (Bakker, 2004; Bright & Hoeffner, 1993); Cai, 1998; Cobb, 2002; Gal, Rothchild & Wagner, 1990; Hughes, 1998; Mokros & Russell, 1995; Shaughnessy, 2006, 1992).

Most of the research in statistics education has been done with primary school children or with university students, resulting in a gap in our knowledge about students' conceptions of averages at the secondary level. According to Begg, Pfannkuch, Camden, Hughes, Noble, & Wild, (2004**),** for many teachers statistics continues to be a content area in which they have little experience since it has only recently become a core area in some curricula. Perhaps more traditional mathematics teaching skills do not transfer into this new domain. This may be due in part to the fact that mathematics is so often taught as a subject focused on procedures. In statistics, surely it is more helpful to place the emphasis on helping students learn to formulate questions, gather data, and use data wisely in solving real problems.

Concerns about the importance of statistics in everyday life and in schools, the lack of research in this area and students' difficulties in statistical reasoning, determined the focus of the larger study (Sharma, 1997). The study investigated the question: What ideas do (Fijian) form five students have about statistics (measures of centre, graphical representations and probability), and how do they construct these? This paper explores student ideas related to averages, just one of the areas explored in the study.

Literature Review

Although the most common statistical idea encountered in everyday life and school contexts is the average value of a set of data, studies of students' understanding of average show that students lack the full range of meanings that "average" is used to convey. This section first draws attention to conceptual and procedural understanding. Then previous research on students' notions of averages is discussed.

Conceptual and Procedural knowledge

Several studies on methods of improving students' general competence appear in reviews by Romberg and Carpenter (1986) and Silver (1990). One of the relevant findings from these studies which helps to extend research on statistical teaching and learning, is the value of connecting concepts and procedures. More time spent on developing understanding leads to increased student performance on problem-solving tasks. This emphasis on understanding is supported by Hawkins, Jolliffe and Glickman, (1992, p. 9) who write that it is "no longer acceptable for students to be taught to perform calculations without understanding what they are doing."

Hiebert and Lindquist (1990, p. 19) define conceptual knowledge as knowledge which “is rich in relationships. It can be thought of as a connected web, where every piece of information is related or connected to other pieces of information.” In contrast, procedural knowledge is made up of rules, procedures or computations for performing mathematical tasks. The researchers assert that instruction should be designed to help students acquire both concepts and procedures.

From another perspective, Rittle-Johnson, Siegler and Alibali’s (2001) teaching experiment about development of conceptual and procedural knowledge of decimal fractions demonstrated that children’s conceptual and procedural knowledge develop iteratively. They defined procedural knowledge as “the ability to execute action sequences to solve problems” and conceptual knowledge as “the implicit or explicit understanding of the principles that govern a domain and of the interrelations between units of knowledge in a domain” (pp. 346-347). Their experiment identified that competence in mathematics requires both conceptual and procedural knowledge. They argue that there is a bi-directional relationship between conceptual understanding and procedural skills. Conceptual understanding is the basis for acquiring procedural skills. Procedural operation can facilitate deeper understanding of concepts. It appears that both conceptual and procedural knowledge are important, so teaching should be designed to help students see connections between them. For instance, if teachers were asked what they would really like students to know one year after completing a topic in statistics, most would respond that they want their students to understand some basic concepts and to be able to apply procedures in meaningful situations. The desirability of making connections between concepts and procedures is not helped by research which deals with them separately. Mokros and Russell (1995) explored mostly students' notions of averages, whereas, the National Assessment of Educational Progress (NAEP) (Brown, Carpenter, Kouba, Lindquist, Silver & Swafford, 1988; Cai, 1998) items primarily dealt with computations in statistics rather than with concepts and understanding. In the present project students' conceptual and procedural knowledge about averages were both explored.

Previous Research on Students’ Understanding of Averages

A number of items on the NAEP (Brown et al., 1988) assessment dealt with students' ability to find and use measures of central tendency and variability for given sets of data. Brown et al. (1988) concluded that although many seventh-grade and eleventh-grade students seem to be able to calculate the mean when asked to do so, the children have a very shallow understanding of the concept of averages. For instance, when asked about the size of the mean of two given numbers relative to those two numbers, only about 40 per cent of the seventh-grade students and about 50% of the eleventh-grade students correctly answered that the average must be halfway between the two numbers.

Subsequent research on the NAEP studies found similar results (Shaughnessy & Zawojewsk, 1999). Six of the seven items administered to students in grades 8 and 12 in the NAEP study dealt with finding and using mean and median (Zawojewski & Heckman, 1997). There was significant growth from 1992 to 1994 in the students’ performance on the item that required them to find the mean and median of particular data sets. However, when asked to select which statistics to use when given data sets, students selected the mean over the median without considering the data distribution or context. The item below (Table 1) required students to understand the context, examine two different data sets and then select and explain their choice of the mean or median. Only 12 percent of students in grade 12 produced responses that were scored as correct or partially correct. From the explanations, it was evident that students believed that the mean was superior to the median. One explanation for this could be curriculum emphasis on the mean in earlier grades. Another contributing factor in students’ frequent selection of the mean over the median may be the design of the question. Since no reason for making the choice was given in the mean or median item, students might have assumed that the statistics were going to be used to compare the attendance figures directly between the two theatres. In fact, reporting identical statistics for each theatre would be useful in making such direct comparisons. Scheaffer (2006) notes that the item could be turned into a good statistics question with a little more context and leeway on the correct answer.

Table 1. Daily attendance at two movie theatres for 5 days and the mean (average) and the median attendance.

	Theater A	Theater B
Day 1	100	72
Day 2	87	97
Day 3	90	70
Day 4	10	71
Day 5	91	100
Mean (average)	75.6	82
Median	90	72

(a) Which statistic, the mean or the median, would you use to describe the typical daily attendance for the 5 days at Theater A? Justify your answer.

(b) Which statistic, the mean or the median, would you use to describe the typical daily attendance for the 5 days at Theater B? Justify your answer.

Cai (1998) examined the relationship between the conceptual understanding of the statistical aspects of the average and computational algorithm (add-them-all-up-and-divide) to determine a missing data value to yield a certain average. Cai (1998) found that while the majority of 250 sixth-grade students knew the averaging algorithm, only half could apply the algorithm to solve contextualized average problems. Many of those who achieved success used strategies and representations such as reversing the averaging algorithm and leveling (pictorial and verbal), while others used guess-and-check approach. Despite the fact that the data were presented in a pictograph, only six percent of the students used the leveling or balancing approaches advocated by other researchers (Friel, 1998; Meyer, Browning, & Channell, 1995). Cai suggests that the concept of average needs to be taught both as a statistical idea for describing, comparing and making sense of data sets and as a computational algorithm for solving contextual problems.

Strauss and Bichler (1988) identified seven properties of the mean and conducted in-depth interviews with Israeli children aged eight to 14 years in order to assess their ability to apply each of the seven properties to real-world problems. The different age groups demonstrated different courses of reasoning to support their understanding of specific properties. Results indicated that children had difficulty in understanding the impact that a zero has on the mean. They tended to think that zero means nothing when in fact it is a legitimate piece of data when calculating the mean. Mevarech (1983) provides evidence that college students mistakenly attributed group structure properties like associativity and closure to the operations of computing means and variances. In particular, the students thought that it was possible to average averages by the add-them-up-and-divide algorithm. Mevarech's results are quite troubling because they occurred after these students had received instruction on descriptive statistics.

Strauss and Bichler's (1988) experiments were repeated (with modifications) in the USA by Leon and Zawojewski (1991). In this study, seven properties of the mean were identified and students' (42 fourth grade students, 61 eighth grade students and 42 college students) understandings of these properties were investigated. Many of the questions were presented in story form and no technical knowledge was required to answer them. The results of this study indicated that performance improved with age and that some properties were more difficult to understand than others. In addition, items in story form were significantly easier to solve than were items presented in numerical format.

Mokros and Russell (1995) used open-ended problems to explore the developing conception of average of students from grades 4, 6 and 8. Since the researchers were interested in students' own preferred strategies about averages, they used tasks that asked students to work backward from a mean to possibilities for a data set that could have that particular mean. In one problem (Potato Chips), students were told that the mean cost of a bag of potato chips was $1.35 and then they were asked to construct a collection of bags that had that mean of $1.35. Mokros and Russell identified and analysed five different constructs of average: average as mode, average as algorithm, average as reasonable, average as midpoint and average as point of balance. These interpretations illustrated ways in which the students were or were not developing useful constructs of averages. Students who thought of average as an algorithm could not make connections from their computational procedures to the actual contextual data. On the potato chip question, some of these students multiplied $1.35 by 9, then divided by 9 again, and every single value in the data set was $1.35.

Watson and Moritz (2000) explored the development of concepts of average in a longitudinal study for students in grades 3 to 9 using interviews. Their questions included: Have you heard of the word average before? What does it mean? They also asked students probing questions from media reports: How do you think they got this average? (three hours of TV per day). Watson and Moritz (2000) also asked their subjects to apply an understanding of the mean to determine total (reverse the averaging process) and to find the average in a weighted mean situation. Student responses were classified in a hierarchical manner that reflected the structure of the task set. They concluded that students conceptual understanding of averages follows a developmental sequence from idiosyncratic responses based on out-of-school experiences, to everyday colloquial terms such as 'normal', to 'most' and 'middle' and finally to the representative nature of 'average'.

Konold and Pollatsek (2002) provide four conceptual interpretations of the mean: mean as typical score, mean as fair-share, mean as a way to reduce data and mean as a signal in noisy processes. According to these authors a signal could be true value with error as noise around it. When an object is weighed repeatedly to determine its actual weight, each observation is viewed as deviating from the actual weight by a measurement error. The average of these scores is interpreted as a close approximation to the true weight. They argue that students need to develop a conceptual understanding of signal and noise in order to understand what an average is and recommend that the mean should be initially introduced to students in relation to comparing data sets.

The research discussed in this section shows that students seem to have little understanding of statistical terms such as the mean and median. While many students could compute means or medians, performance was much lower on items that required a deeper understanding of the concept of the average. It seems that students understood the average as an abstract formula, devoid of any conceptual meaning. Konold and Pollatsek (2002) think that part of this problem is due to the interpretations teachers use to introduce data summaries. They suggest that, since thinking about averages as fair shares or typical is not very helpful in developing an aggregate view, such thinking should not be emphasized with students. However, Shaughnessy (2006) argues that mean as fair-share and mean as typical value are a better first introduction to the notions of averages because they build on students' primary intuitions.

Overview of the Study

The larger study (Sharma, 1997) was designed to investigate the question: What ideas do form five students have about statistics (measures of centre, graphical representations and probability), and how do they construct these? The study took place in a co-educational, private secondary school in Fiji. The class consisted of 29 students aged 14 to 16 years. According to the teacher, none of the students in the sample had previously received any in-depth instruction in statistics. Fourteen students were chosen from the class, the criteria for selection including gender and achievement. While the students had been taught how to compute an average as part of their regular mathematics class, they had had no exposure to collecting data and drawing conclusions.

Tasks

To explore the full range of students’ thinking about averages, both open-ended and closed questions were selected and adapted from those used by other researchers (Watson & Callingham, 2003), The weight (Item 1A) and pocket money (Item 1B) questions were used to explore students' procedural understanding of the mean and median.

Item 1A: Weight problem

A small object was weighed on the same set of scales separately by nine students in a science class. The weights (in grams) recorded by each student were: 6.3 6.0 6.0 15.3 6.1 6.3 6.2 6.15 6.3. How can you find the mean weight? What is the mean weight? What is the median weight?

Item 1B: Pocket money problem

Five students' pocket money is: \$1, \$2, \$2, \$3, \$20. What is the mean? What is the median? Which is the more appropriate measure to use in this distribution? Why?

Two questions were used to examine students' knowledge about averages beyond its algorithmic use in mathematics classes. These were about averages in the domain of weather forecasting (Item 2). It was thought that this domain exposes students to collections of events and to attempts at summarizing and making sense of information or to comparing data sets. The first question was designed to elicit the range of conceptual definitions that students have for the word ‘average’. The second was designed to explore students' awareness of the use of averages and the purposes for data reduction summary representations.

Item 2: Weather problem

Sometimes on the radio the weather-person says something like the average temperature this week was 55 degrees. (i) What does the phrase average temperature mean? How does the weather person find the average temperature? (ii) If the weather-person knows the temperature on each of the days of the week, why does s/he want to know the average temperature for the whole week?

Two tasks were used to explore whether students could apply average ideas to more complex situations. The age problem (Item 3A) required students to work backward and find the total when given the mean and then to find the average age of 5 girls. The frequency distribution table of marks (Item 3B) examined how students think about the relationship between measures of centres and data sets presented in tables.

Item 3A: Age problem

The mean age of four girls is 14 years. If a fifth girl has an age of 16, what is the total age of the five girls? What is the mean age?

Item 3B: Mathematics marks of 15 students in a class are:

Mark (x)	**Frequency (f)**
2	3
4	4
6	2
8	1
10	5

How can you find the mean mark? What is the median mark?

Interviews

Each student was interviewed individually, by the researcher, in a room away from the rest of the class. The interviews were tape recorded for analysis, and notes were made of student non – verbal behaviours observed during the interview. Each interview lasted about 40 to 50 minutes. Paper, pencil and a calculator were provided for the student if he or she needed it.

Analysis and coding of data

The data revealed that many of the students applied rules and procedures inappropriately or used strategies based on prior everyday and school experiences. I created a simple four category rubric that could be helpful for describing research results relating to students' statistical conceptions, planning instruction in statistics and dissemination of findings to mathematics educators. The four categories in the model are: non-response, non-statistical, partial-statistical and statistical. The term statistical is used in this article for the appropriate responses. However, I am aware that such a term is not an absolute one. Students possess interpretations and representations which may be situation-specific, and hence these ideas have to be considered in their own right. This category has been used mainly to discuss and present results. It would be reasonable to assume another level (advanced-statistical), equivalent to Shaughnessy's (1992) pragmatical statistical level, where students appear to have a very complete view that incorporates questioning of data, but the need for such a category did not arise in my research and any responses that could have been categorised as advanced-statistical were simply grouped with the statistical responses. The main focus in this article is on the non-statistical responses (in which students made inappropriate connections with everyday experiences) and the partial-statistical responses (in which students applied rules and procedures inappropriately). Extracts from typical individual interviews are used for illustrative purposes. Throughout the discussion, *I* is used for the interviewer and *Sn* for the nth student

Results

Procedures for finding the mean

As indicated earlier, the weight (Item 1A) and pocket money (Item 1B) tasks were used to explore students' procedural understanding of the mean. Results are summarised in Table 2.

Table 2. Response types for tasks involving procedures for finding the mean

(n = 14)

Response type	**Number of students using it**		
	Weight task (Item 1A)	PM task (Item 1B)	[Both tasks] *
Non-response	3	1	[1]
Non-statistical	–	–	–
Partial-statistical	3	1	–
Statistical	8	12	[8]

* In this and subsequent tables, the figures in the [both tasks] column refer to the number of students whose response was the same on both tasks.

Table 2 reveals that a small minority of students had difficulty responding at all and the majority responded in a statistical fashion. To be considered statistical, students had to add the data values and divide the totals by the number of values. Students who applied the averaging algorithm but made minor calculation errors were coded as statistical. A possible explanation of why more students were statistical on the pocket money problem than on the weight problem is that the money context had been used for calculating the mean in class and students use money on a daily basis. The teacher had asked five students how much money they had in their pockets and listed the amounts on the board. The students calculated the mean from the data tabulated by the teacher. Another explanation for this could be the type of numbers and rounding involved in the two problems. Decimal numbers certainly created problems for some students although a calculator was available if they wanted to use it. It must be noted that none of the students commented on the weight of 15.3g to be investigated. Perhaps it was a mistake.

Partial statistical responses: Although only four students replied in a partial-statistical way, their responses provide an insight into the effect of relying on rules. The students remembered various rules and procedures but applied them inappropriately, working out the median or range instead of the mean.

Conceptual ideas about the mean

In contrast to the procedural responses, responses this time indicated wide differences in level of sophistication and quality of explanations for the questions: ‘What is an average?’ and ‘What is it used for?’ The results are summarised in Table 3.

Non-responses and statistical responses were approximately equivalent in number and comprised the minority. An additional non-response strategy emerged in relation to this weather task. Part of the question was sometimes repeated. When asked what a weather person meant by the statement ‘the average temperature this week was 55 degrees’, student 3 said that, “it was average for the week.” The three students whose reasoning was classified as statistical (Item 2i) could clearly explain that the average temperature was the mean temperature and that it was determined by adding the temperatures for the seven days and dividing by the number of days (7). In two cases (Item 2ii), student responses were coded as statistical. The responses suggest a data reduction and summary representation view of averages. For instance, one student said that the weather person needs the mean so that he does not have to go over all the temperatures.

Table 3. Response types for tasks involving mean concepts (n = 14)

Response type	Number of students using it		
	Weather item (2i)	Weather item (2ii)	[Both tasks]
Non-response	3	3	[2]
Non-statistical	3	6	[3]
Partial-statistical	5	3	[2]
Statistical	3	2	[2]

Non-statistical responses: Three students used non-statistical responses for the first task and six for the second. Their responses were based upon everyday (cultural) and school experiences which often involve the use of language in non-mathematical ways. The students referred to average temperature as the 'usual 'or 'ordinary' temperature recorded each week. Student 21 based his reasoning on his prior school experiences, referring to average temperature as 'border-line temperature', possibly because at times teachers refer to students sitting on the average mark in test results as borderline cases.

Similarly, in response to the question of why the weather person would want to find the average temperature, these students did not mention the summary statistics aspect but drew upon their cultural experiences, talked about the announcer's job, or how people benefit from weather reports, or how to decide which day is cloudy or rainy. Further probing of these students' ideas failed to elicit any deeper reasoning. Watson and Moritz (2000) chose to categorise such responses at the pre-average level.

Partial-statistical responses: The students whose responses were classified as partial-statistical on both the tasks had only a procedural understanding of the mean. This is illustrated by student 6 who, when asked what a weather person meant by the statement "the average temperature this week was 55 degrees", responded:

S6: 7.85.
I: How did you work it out?
S6: 55 degrees divide by 7.
I: Why did you divide by 7?
S6: For the week. Each day the temperature was 55 degrees. For the week it was 7.85.

The other students, whose responses were in the partial-statistical category, mentioned partial properties of average. They explained that the average was the mean temperature and asserted that it was determined by the sum of scores divided by the number of scores. When asked what was added and divided in the weather problem, the students could go no further.

Applications of the mean

Two questions were used to explore whether students could apply mean ideas. One was an age problem (Item 3A), and the other a frequency distribution table of marks (Item 3B). Results are summarised in Table 4.

Table 4. Response types for tasks involving age problem and frequency tables (n = 14)

Response type	**Number of students using it**		
	Age task (Item 3A)	Frequency task (Item 3B)	[Both tasks]
Non-response	1	2	[1]
Non-statistical	–	–	–
Partial-statistical	6	5	[5]
Statistical	7	7	[7]

The results in Table 4 bear a resemblance to those in Table 2. There were very few non-responses and a substantial number of statistical replies. The latter group did not rely completely on the add-them-and-divide algorithm. They saw the need for the average to be weighted by sample size.

Partial-statistical responses: Five students applied the mean formula inappropriately to both questions. When asked for the total age and mean age of five girls, given that the mean age of 4 girls is 14 years and a 5th girl has an age of 16, all five students said that the total age was 30 (14 + 16) and the mean age was 6 (30/5). The task, on the other hand, required them to compute 4 times 14, plus 16, divided by 5.

When asked to consider the mean from the frequency table (Item 7.4), the students either adapted or applied the mean rules inappropriately. Three of these students simply added the marks column and divided the result by the total frequency (30/15) to get a mean of 2. Two students divided the total mark (30) by the highest frequency (5) and obtained a result of 6.

Procedures for finding the median

The median, as a measure of central tendency of data, is the number which has half the values above it and half below it when the values are arranged in order of size. It may be the central value or it may be the average of the middle two. The weight (Item 1A) and the pocket money tasks (Item 1B) were also used to explore students' procedural understanding about the median. The students' responses to the two questions are summarised in Table 5.

Table 5. Response types for tasks involving procedures for finding the median (n = 14)

Response type	**Number of students using it**		
	Weight task (Item 1A)	PM task (Item 1B)	[Both tasks]
Non-response	3	1	[1]
Non-statistical	1	–	–
Partial-statistical	5	1	[1]
Statistical	5	12	[5]

The results on these median tasks are similar to those on the same tasks involving the mean (Table 1). A minority of replies were considered non-responses, while a majority represented statistical thinking. The finding that just over twice as many statistical responses occurred on the pocket money problem than on the weight problem may have resulted from a difference in the arrangement of the data. Since the weight data were not given in any order, the students had to first arrange the data in order, which caused difficulty for some. The money problem data did not require any re-arrangement and hence caused less difficulty.

Non-statistical responses: Only one student response was classified as non-statistical but the response was interesting because the student (Student 22) seemed to make inappropriate use of previous learning about sets.

I: What is the median value of this set [weight] of data?
S22: Middle score. Arrange in ascending order and then find the middle.
I: How did you get 6?
S22: There are six scores.
I: But I think there are nine scores.
S22: But when we arrange in order, we should not repeat an element.

This knowledge about sets was most likely gained from studying Book 3A (Fijian Ministry of Education, 1977a) in which there is a unit on *Sets*. In this unit students learn that one does not repeat the elements when listing a set.

Partial-statistical responses: Responses classified as partial-statistical employed two strategies that teachers need to be aware of. Three of the five students who gave partial-statistical responses on the weight problem, applied mean, mode and range rules inappropriately to the median task. Even when questioned as to whether the *same* procedures were appropriate for finding both the mean and median, the students who had applied the mean procedures for this task did not change their answer. The other student and the student who gave partial-statistical response on both the tasks adapted the median rule. They did not arrange the weights in order but simply found the middle weight.

Conceptual ideas about the median
Two tasks were used to explore students' conceptual understanding of median, namely providing a definition of median and the pocket money problem (Item 1B). Their responses are summarised in Table 6.

Table 6. Response types for tasks involving median concepts (n = 14)

Response type	**Number of students using it**		
	Definition task	PM task (Item 1B)	[Both questions]
Non-response	1	4	[1]
Non-statistical	1	6	–
Partial-statistical	9	2	[2]
Statistical	3	2	[2]

Table 6 data reveal that a minority of students displayed non-response or statistical reasoning with respect to the two tasks. The students whose thinking was classified as statistical not only knew how to calculate the median, but could also provide appropriate meanings for the term median and state the disadvantages of using the mean. Further the students gave a reason to explain, although with something of a struggle, when it is appropriate to use the mean and when it is appropriate to use the median. That is, the median is used mainly when extreme values will unduly affect the mean, making the mean less representative of the data set as a whole. The majority of student responses (18 out of 28) fell into the middle two categories of non-statistical and partial-statistical, the two which are most revealing of students' unusual ideas.

Non-statistical responses: Prior school experiences and knowledge seemed to lie behind several of the non-statistical responses. For example, in constructing a meaning for median, one student referred to the median as the middle or the centre. When asked to explain middle or centre of what, the student talked about numbers and circles. The student appeared to relate his geometry knowledge to this statistics question.

When asked to choose the more appropriate measure to use, the mean or the median, the students believed that mean was the more appropriate to use because everyone will get equal amounts. They did not realise that in this particular example the mean was affected by the extreme value ($20) and hence did not give a good picture of the distribution. Student 5 was typical of this group.

I: Which one do you think is a more appropriate measure to use?
S5: Mean.
I: Why?
S5: If we find the mean, the person with $20 will get equal amount.

Partial-statistical responses: The students who gave partial-statistical responses had only procedural understanding about median, or understood the median was somehow connected with middle value but could not explain middle of what. Student 17 was one of these.

S17: Arrange the numbers in order and then find the median value.

On the pocket money item, these students said that the median was the more appropriate measure to use. However, when asked to explain their reasoning they said that it was the middle value. Even further probing did not elicit any idea about the mean being affected by the extreme value.

Discussion

Overall, the results show that most students in the study had only a procedural understanding of mean and median. While students could work out the summary statistics involved in simple data sets (Items 1A, 1B), they applied the mean and median algorithmic procedures inappropriately to the questions which involved frequency table (Item 3B). Although the teacher had taught 'finding the averages' in a number of ways, including from a frequency distribution table, the students appeared to have muddled views as to how to apply these procedures. When applying the median concept to a frequency distribution (Item 3B), students chose the middle frequency or the largest frequency instead of the appropriate observation. Additionally, it was difficult for students to remember which rules went with which problem. Only a very small number of students knew how to calculate averages, provide appropriate meanings for the terms, and state the advantages and disadvantages of using the measure.

Some students even thought that it is possible to average averages by the add-them-up-and-divide algorithm. It appears that the limited nature of school instruction in statistics did not provide them with opportunities to refine their concepts and resolve ambiguities. The finding that most students have only a procedural understanding of summary statistics, is consistent with the results of American studies discussed earlier (Brown et al., 1988; Gal et al., 1990; Konold & Pollatsek, 2002; Pollatsek et al., 1981; Russell and Mokros, 1995; Shaughnessy & Zawojewsk, 1999). Pollatsek et al. reported that even college students understood the average as an abstract formula, devoid of any conceptual meaning. It appears that most Fijian students lacked a general understanding of the concepts involved, although some constructed an understanding of the concepts from the context.

The results show that students entertained several ideas about the concepts of mean and median, some of which were based on what they had absorbed formally or informally from their everyday environments. The finding that students base their thinking in statistics on their prior experiences is not new. Watson and Moritz (2000) report that the thinking of many students in their study was dominated by past experience and this prevented them developing statistical ideas, despite being aware of notions of averages. In some respects, the findings of the present investigation go beyond those discussed in research literature. The findings demonstrate how students' other school experiences also influence their construction of statistical ideas. At times the in-school experiences appear to have had a negative effect on the students. An example of negative effect that arose from other school experiences were the students who were deeply convinced that one cannot repeat elements in a data set. On the other hand, if meanings and understandings associated with averages are tied to the specific situations (Lave, 1991), and structured by social situations, student opinions cannot be judged as incorrect (Gal, 1998). Perhaps, it is important to point out to students that there are alternative points of view.

The results of the interviews show that although contexts may help students use prior knowledge, such situational knowledge is diverse and can also cause misinterpretations of the information in the data display. Student 21's personalisation of the context brought in various interpretations of the task (Item 2) and inconsistency in his explanations. Probably, the student's lack of understanding of the constraints imposed by the context distracted him from making a sensible interpretation. Given that statistics is often taught through examples drawn from "real life", teachers need to exercise care in ensuring that the intended support apparatus is not counterproductive. This is particularly important in light of current curricula calls for pervasive use of contexts (Ministry of Education, 2007; Rossman, Chance and Medina, 2006; Scheaffer. 2006; Watson, 2004) and research showing the effects of contexts on students' ability to solve open-ended tasks (Cooper and Dunne, 1997; Lubienski, 2007; Meyer, Dekker, & Querelle, 2001; Sullivan, Zevenbergen & Mousley, 2002). The study by Cooper and Dunne showed that realistic problems disadvantaged working class children since middle class children had greater linguistic facility.

Conversely, in spite of the importance of relating classroom mathematics to the real world, the results of my research indicate that students frequently fail to connect the mathematics they learn at school with situations in which it is needed. While students could calculate summary statistics from simple data sets (Item1B), they had difficulty explaining why the weather person needs to know the average temperature (Item 2). Clearly, the results support claims made by other researchers (Gal et al., 1990; Konold, & Pollatsek, 2002; Mokros & Russell, 1995). Konold and Pollatsek claim that students rarely use the mean in the context of comparing data sets. Mokros and Russell state that the type of conceptions that emerge may depend on tasks used. It appears that learning for these students is situation specific and that connecting students' everyday contexts to academic mathematics in a way that enhances meaning, is not easy.

Konold, and Pollatsek (2002) suggest that viewing averages as fair share may not provide an appropriate conceptual basis for representing an entire group. The findings of my study seem to support this claim. Some students viewed averages as fair shares. Student 5 used the mean rather than the median (Item 1B) because "everyone will get equal amount." The student did not think of the computed value in relation to other individual values or whether there was an outlier.

Students' overall reluctance to use the median is consistent with the findings of (Zawojewski & Heckman, 1997). From students' explanations, it was evident that students believed that the mean was superior to the median. Another contributing factor in students' frequent selection of the mean over the median may be the design of the question. Since no reason for making the choice was given in Item 1B, students might have assumed that the statistics was going to be used to compare the pocket money figures directly, so they chose the mean.

Implications for Teaching and Research

This paper has focused on students' responses to mean and median tasks. The non-statistical responses point to some inappropriate connections being made with other parts of mathematics. The partial-statistical responses point to some misconceptions and alternative ideas which led students to using rules, but often inappropriately. These issues need to be addressed in high school mathematics courses to ensure that students develop both the conceptual and procedural understanding of the concept of average.

One implication for further research could be to replicate the present study and include a larger sample of students from different ethnic and educational backgrounds to claim generality. Since interviewing can be time consuming, probably there is a need to conduct a survey with these different groups. A sample of these students could be interviewed in order to probe their conceptions of averages at a greater depth.

Another implication relates to meaningful contexts. The picture of students' thinking in regards to averages is somehow limited because students responded to few items. There is a need to include more items using different contexts such as comparing data sets presented in tables and graphs, and applications to problem-solving settings, in order to explore students' conceptions of averages and related contexts in much more depth. Extending the question to include the reversing of the averaging process (Cai, 1998) and Watson and Moritz's (1999) graphical representation tasks might also be used.

Research (Russell & Mokros, 1995) shows that task context can bring in multiple interpretations and possibly different kinds of abstractions. While 12 students managed to respond in a statistical way on Item 1B (Table 2), only eight responses were considered appropriate on Item 1A (Table 2). At this point it is not clear how a learner's understanding of the context contributes to his/her interpretation of data. Research on what makes this connection difficult for students is needed.

Fourth, this small scale investigation into identifying and describing students' reasoning in regards to the concept of averages has opened up possibilities for further research at a macro-level on students' thinking and to develop more explicit categories for each level of Shaughnessy's (1992) framework. Such research would validate the framework of response levels described in the current study and raise more awareness of the levels of thinking that need to be considered when planning instruction and developing students' statistical thinking.

Many students were part-way to providing a complete explanation, but needed more detail or precision. Teachers need to assist students to express what they already know about averages with more precise statistical language. In the course of discussions, comparison of several interpretations of the word *average* may be made. This may lead to judgments about what might constitute a good explanation, draw attention to missing details and help students develop written and oral communication skills.

Teachers need to explicitly include activities that promote conceptual understanding in statistical analyses in statistics education. While it is common practice in classrooms to show students how to compute mean and median from supplied data, students are seldom asked to collect data and then choose an appropriate measure which is part of a meaningful context. Using manipulative tools such as cuisenaire rods (Hollingsworth, 1995), can offer a path for connecting conceptual and procedural understanding about averages.

Like the students, teachers may resort to partial-statistical or non-statistical explanations. Research efforts at the in-service and pre-service levels are crucial in order to better understand how teachers view measures of central tendency and to inform teacher educators and curriculum writers. Responses to average tasks, such as those used in research (Watson & Callingham, 2003; Watson & Moritz, 1999), could provide starting points for exploring the concepts of averages.

Ideally, it would be good to make links between students' responses and what happens in classrooms. This could be achieved by doing classroom observations, as well as by collecting data from teachers. This could enable researchers to gain insights into why the students applied rules and procedures inappropriately. From lesson observations, researchers could collect examples of best practice for dissemination amongst the wider mathematics education community.

References

Australian Education Council. (1991). *A national statement on mathematics for Australian schools*. Canberra: Curriculum Corporation.

Bakker, A. (2004). Reasoning about shape as a pattern in variability. *Statistics Education Research Journal,* 3(2), 64-83.

Begg, A., Pfannkuch, M., Camden, M. Hughes, P., Noble, A. & Wild, C. (2004). *The school statistics curriculum: Statistics and probability education literature review.* Auckland: Auckland Uniservices Ltd, University of Auckland.

Bright, G. B. & Hoeffner, K. (1993). Measurement, probability, statistics and graphing. In T. D. Owens (Ed.), *Research ideas for the classroom: Middle grades mathematics* (pp. 78-98). New York: Macmillan.

Brown, C. A., Carpenter, T. P., Kouba, V.L., Lindquist, M. M. Silver, E. A. & Swafford, J. O. (1988). Secondary school results for the fourth NAEP mathematics assessment: Discrete mathematics, data organisation and interpretation, measurement, number and operations. *The Mathematics Teacher,* 81(4), 241-248.

Cai, J. (1998). Exploring students' conceptual understanding of the averaging algorithm. *School Science and Mathematics,* 98, 93-98.

Cobb, P. (2002). Modeling, symbolizing, and tool use in statistical data analysis. In K. Gravemeijer, R. Lehrer, B. V. Oers & L. Verschaffel (Eds.), *Symbolizing, modeling and tool use in mathematics education* (pp. 171-195). Dordrecht: Kluwer Academic Publishers.

Cooper, B. & Dunne, M. (1999). *Assessing children's mathematical ability*. Buckingham: Open University Press.

Fijian Ministry of Education, Women, Culture, Science and Technology (1994). *Primary mathematics prescriptions.* Suva: Curriculum Development Unit.

Fijian Ministry of Education. (1977a). *Mathematics book 3A.* Suva: Curriculum Development Unit.

Friel, S. N. (1998). Teaching statistics: What's average? In L. J. Morrow (Ed.), *The teaching and learning of algorithms in school mathematics* (pp. 208-217). Reston, VA: National Council of Teachers of Mathematics.

Gal, I. (1998). Assessing statistical knowledge as it relates to students' interpretation of data. In S. P. Lajoie (Ed.), *Reflections on statistics: Learning, teaching, and assessment in grades K-12* (pp. 275-295). *Mahwah* NJ: Lawrence Erlbaum Associates.

Gal, I. (1995). Statistical tools and statistical literacy: The case of the average. *Teaching Statistics,* 17, 97-99.

Gal, I., Rothchild, K. & Wagner, D. A. (1990). *Statistical concepts and statistical reasoning in school children: Convergence or divergence.* Paper presented at the American Educational Research Association Conference, Boston.

Hawkins, A., Jolliffe, F. & Glickman, L. (1992). *Teaching statistical concepts.* New York: Longman Publishing.

Hiebert, J. & Lindquist, M. (1990). Developing mathematical knowledge in the young child. In N. Payne (Ed) *Mathematics for the young child,* (pp. 17-38). Reston, VA: National Council of Teachers of Mathematics.

Holmes, P. (1994). Teaching statistics at school level in some European countries. In L. Brunelli and G. Cicchutelli (Eds.), *Proceedings of the first scientific meeting* (pp. 3-11). University of Perugia, 23-24 August 1993.

Hollinsworth, C. (1995). Beginning statistics with Cuisenaire rods. *Teaching Statistics,* 17(2), 38-39.

Hughes, P. (1998). Constructing meaning for some statistical measures. *The New Zealand Mathematics Magazine,* 35(2), 8-14.

Konold, C. & Pollatsek, A. (2002). Data analysis as the search for signals in noisy processes. *Journal for Research in Mathematics Education,* 33(4), 259-289.

Konold, C., Well, A., Lohmeier J. & Pollatsek, A. (1993). *Understanding the law of large numbers.* Paper presented at the Fifteenth Annual Conference of the North American Chapter of the International Group for the Psychology of Mathematics Education, Pacific Grove, CA, October, 1993.

Lave, J. (1991). *Cognition in practice: Mind, mathematics and culture in everyday life*. New York: Cambridge University Press.

Leon, M. R., & Zawojewski, J. S. (1991). Use of the arithmetic mean: An investigation of four properties, issues and preliminary results. In D. Vere-Jones (Ed.), *Proceedings of the Third International Conference on Teaching Statistics: Vol. 1. School and general issues* (pp. 302-306). Voorburg, The Netherlands: International Statistical Institute.

Lubienski, S. T. (2007). Research, reform, and equity in U.S. mathematics education. In N. S. Nasir & P. Cobb (Eds.), *Improving access to mathematics: Diversity and equity in the classroom* (pp.10-23). New York: Teachers College Press.

Mevarech, Z. R. (1983). A deep structure model of students' statistical misconceptions, *Educational Studies in Mathematics*, 14, 415-429.

Meyer, R. A., Browning, C., & Channell, D. (1995). Expanding students' conceptions of the arithmetic mean. *School Science and Mathematics,* 95, 114-117.

Meyer, M. R, Dekker, T. & Querelle, N. (2001). Context in mathematics curricula. *Mathematics Teaching in the Middle School,* 6(9), 522-527.

Ministry of Education. (2007). *The New Zealand curriculum*. Wellington: Ministry of Education.

Ministry of Education. (1992). *Mathematics in the New Zealand Curriculum.* Wellington: Ministry of Education.

Mokros, J. & Russell, S. J. (1995). Children’s concepts of average and representativeness. *Journal for Research in Mathematics Education*, 26, 20-39.

Pollatsek, A., Lima, S. & Well, D. (1981). Concept or computation: Students' understanding of the mean. *Educational Studies in Mathematics*, 12, 191-204.

Rittle-Johnson, B., Siegler, R. & Alibali, M. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology, 93*(2), 346-362.

Romberg, T. A. & Carpenter, T. (1986). Research on teaching and learning mathematics: Two disciplines of scientific inquiry. In M. Wittrock (Ed.), *Handbook of research in teaching* (pp. 850-873). New York: Macmillan.

Rossman, A., Chance, B. & Medina, E. (2006). Some important comparisons between statistics and mathematics, and why teachers should care. In G. Burrill & P. C. Elliot (Eds.), *Thinking and reasoning with data and chance* (pp. 323-333). Reston: The National Council of Teachers of Mathematics.

Scheaffer, R. L. (2006). Statistics and mathematics: On making a happy marriage. In G. Burrill & P. C. Elliot (Eds.), *Thinking and reasoning with data and chance* (pp. 309-322). Reston: The National Council of Teachers of Mathematics.

Scheaffer, R. L, Watkins, A. E. & Landwehr, J. M. (1998). What every high school graduate should know about statistics? In S. P. Lajoie (Ed.), *Reflections on statistics: Learning, teaching, and assessment in grades K-12* (pp. 3-27). Mahwah, NJ: Erlbaum.

Sharma, S. (1997). *Statistical ideas of high school students: Some findings from Fiji.* Unpublished doctoral thesis. Waikato University, Hamilton, New Zealand.

Shaughnessy, J. M. (2006). Research on students' understanding of some big concepts in statistics. In G. Burrill & P. C. Elliot (Eds.), *Thinking and reasoning with data and chance* (pp. 77-78). Reston: The National Council of Teachers of Mathematics of Teachers of Mathematics.

Shaughnessy, J. M. (1992). Research in probability and statistics: Reflections and directions. In D. Grouws (Ed.). *Handbook of research on mathematics teaching and learning* (pp. 309-321). New York: Macmillan.

Shaughnessy, J. & Pfannkuch, M. (2004). How faithful is old faithful? Statistical thinking: A story of variation and prediction. *Mathematics Teacher*, 95(4), 252-259.

Shaughnessy, J. M. & Zawojewski, J. S. (1999). Secondary students' performance on data and chance in the 1996 NAEP. *The Mathematics Teacher*, 92(8), 713-718.

Silver, E. (1990). Contributions to research to practice: Applying findings, methods and perspectives. In T. Cooney (Ed.), *Teaching and learning mathematics in the 1990s* (pp. 1-11). Reston, VA: National Council of Teachers of Mathematics.

Strauss, S. & Bichler, E. (1988). The development of children's concepts of the arithmetic average. *Journal for Research in Mathematics Education*, 19(1), 74-80.

Sullivan, P., Zevenbergen, R., & Mousley, J. (2002). Contexts in mathematics teaching: Snakes or ladders? In B. Barton, K. C. Irwin, M. Pfannncuch & M. Thomas (Eds.), *Mathematics education in the South Pacific* (pp. 649-656). Proceedings of the 25th annual conference of the Mathematics Education Research Group of Australasia, Auckland.

Wallman, K. K. (1993). Enhancing statistical literacy: Enriching our society. *Journal of the American Statistical Association*, 88(421), 1-8.

Watson, J. M. (2004). Quantitative literacy in the media: An arena for problem solving. *Australian Mathematics Teacher,* 60(1), 34-40.

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

Watson, J. M. & Moritz, J. B. (2000). The longitudinal development of understanding of average. *Mathematical Thinking and Learning* 2(1), 9-48.

Watson, J. M. & Moritz, J. B. (1999). The beginning of statistical inference: Comparing two data sets. *Educational Studies in Mathematics*, 37, 145-168.

Zawojewski, J. S. & Heckman, D. S. (1997). What do students know about data analysis, statistics, and probability? In P. A. Kenney and E. A. Silver (Eds.), *Results from the sixth mathematics assessment of the National Assessment of Educational Progress* (pp. 1-15). Reston VA: The National Council of Teachers of Mathematics of Teachers of Mathematics.