TIM JACOBBE

ELEMENTARY SCHOOL TEACHERS' UNDERSTANDING OF THE MEAN AND MEDIAN

Received: 25 March 2010; Accepted: 29 September 2011

ABSTRACT. This study provides a snapshot of elementary school teachers' understanding of the mean and median. The research is presented in light of recent work regarding preservice teachers' understanding of the mean. Common misconceptions are identified which lead to potential implications for teacher preparation programs. One of the primary concerns regarding increasing the standards expected of students to learn statistics is teachers' preparation to address those standards. Exploring issues with teachers' understanding of two of the most prominent concepts in the enacted curriculum provides a glimpse into the need to adequately prepare teachers to teach statistics.

KEY WORDS: content knowledge, mean, median, statistics education, teachers' preparation

Introduction

In the USA, the efforts of the Quantitative Literacy Project (Scheaffer, 1986) led the National Council of Teachers of Mathematics (NCTM) to gradually increase the depth of statistical topics covered in elementary, middle, and secondary schools (NCTM, 2000). In order to support the objectives set forth in the *Principles and Standards for School Mathematics*, the American Statistical Association released a curriculum framework for PreK–12 statistics education entitled Guidelines for Assessment and Instruction of Statistics Education (GAISE) (Franklin, Kader, Mewborn, Moreno, Peck, Perry & Scheaffer, 2007). As defined by the authors of the GAISE, "This framework provides a conceptual structure for statistics education which gives a coherent picture of the overall curriculum. This structure adds to but does not replace the NCTM recommendations" (Franklin et al., 2007, p. 5). In light of recent reform movements in the USA to include more sophisticated statistical topics in the K–12 setting, the study presented in this paper explored three elementary school teachers' understanding of the mean and median.

One of the primary concerns that motivated the creation of the GAISE framework was that "statistics…is a relatively new subject for many teachers who have not had an opportunity to develop sound understand-

International Journal of Science and Mathematics Education (2012) 10: 1143–1161

ing of the principles and concepts underlying the practices of data analysis they are now called upon to teach" (Franklin et al., 2007, p. 5). With these expansions to the K–12 curriculum, it is important to examine what teachers know about the subject matter. The mean and median are considered for the purposes of this paper since these topics have been present in elementary schools in the USA for over 100 years (Watson, Callingham & Kelly, 2007). If elementary school teachers have difficulty with these topics, then caution should be used when increasing the level of sophistication at which statistics is covered in schools without addressing teachers' preparation to teach statistics effectively.

In regard to teaching statistics, the issue of teachers' preparation is also raised by Shaughnessy when he comments that "teachers' backgrounds are weak or nonexistent in stochastics and in problem solving. This is not their fault, as historically our teacher preparation programs have not systematically included either stochastics or problem solving for prospective mathematics teachers" (1992, p. 467). This concern is reiterated by Shaughnessy in his most recent chapter in the *Second Handbook of Research on Mathematics Teaching and Learning* when he states, "Most K–12 teachers in the United States have very little background in statistics" (2007, p. 995). It should be clear that the results presented in this paper are not provided as a means to criticize the knowledge of teachers; rather, they are provided as an attempt to provide a snapshot of what may be expected in a system that does not prepare teachers to teach statistics. Particularly in the elementary school, "the prevailing assumption is that the content of the K–12 curriculum is already understood by teachers and …is relatively simple. However, in reality, what teachers have learned about mathematics in their pre-college mathematics classes is not adequate for teaching mathematics [or statistics] for understanding" (NCTM, 1991, p. 74).

The GAISE Framework and the Mean and Median

The authors of the GAISE framework identify three levels of statistical development (levels A, B, and C) that students must progress through in order to develop statistical understanding. It is paramount for students to have worthwhile experiences at level A during their elementary school years in order to prepare for future development at levels B and C at the middle and secondary levels. "Without such experiences, a middle (or high) school student who has had no prior experience with statistics will need to begin with Level A concepts and activities before moving to

Level B" (Franklin et al., 2007, p. 13). The GAISE framework indicates that students at level A should understand the mean as a fair share and the median as the middle point.

The measures of central tendency introduced at level A are also expanded upon at level B. The biggest expansion to the measures of center introduced at level A is that students should begin to see the mean as a "balance point" rather than as a "fair share" (Franklin et al., 2007, p. 41). The following activity provided by the GAISE gives an example of how students should visualize this concept.

Nine students were asked: How many pets do you have? The resulting data were 1, 3, 4, 4, 4, 5, 7, 8, 9. [These data are summarized in a dotplot]

If the pets are combined into one group, there are a total of 45 pets.

If the pets are evenly redistributed among the nine students, then each student would get five pets. That is, the mean number of pets is five. [The dotplot is then presented with 9 dots above the 5]

It is hopefully obvious that if a pivot is placed at the value 5, then the horizontal axis will 'balance' at this pivot point. That is, the 'balance point' for the horizontal axis for this dotplot is 5. What is the balance point for the dotplot displaying the original data? We begin by noting what happens if one of the dots over 5 is removed and placed over the value 7 [They show a dotplot with 8 dots over 5 and one dot over 7]. Clearly, if the pivot remains at 5, the horizontal axis will tilt to the right. What can be done to the remaining dots over 5 to 'rebalance' the horizontal axis at the pivot point? Since 7 is two units above 5, one solution is to move a dot two units below 5–3, as shown below [A dotplot is shown with 1 dot over 3, 7 dots over 5, and 1 dot over 7].

The horizontal axis is now rebalanced at the pivot point. Is the only way to rebalance the axis at 5? No. Another way to rebalance the axis at the pivot point would be to move two dots from 5 to 4, as shown below [A dotplot is shown with 2 dots above 4, 7 above 5, and 1 above 7].

The horizontal axis is now rebalanced at the pivot point. That is, the 'balance point' for the horizontal axis for this dotplot is 5. Replacing each dot in this plot with the distance between the value and 5 we have [There is a dotplot with dots replaced by the distance away from 5, so there are two 1's above 4, seven 0's above 5, and one 2 above 7]. Notice that the total distance for the two values below the 5 (the two 4's) is the same as the total distance for the one value above the 5 (the 7). For this reason, the balance point of the horizontal axis is 5. Replacing each value in the dotplot of the original data by its distance from 5 yields the following plot [There is a dotplot with one 4 above 1, one 2 above 3, three 1's above 4, one 0 above 5, one 2 above 7, one 3 above 8, and one 4 above 9].

The total distance for the values below 5 is 9, the same as the total distance for the values above 5. For this reason, the mean (5) is the balance point of the horizontal axis.

Franklin et al., 2007, pp. 41–43

Theoretical Perspective

Over the past 20 years, several studies have been conducted on students' understanding of average (e.g. Batanero, Cobo & Diaz, 2003; Cai & Moyer, 1995; Garcia & Garret, 2006; McGatha, Cobb & McClain, 1998; Russell & Mokros, 1991; Watson & Moritz, 1999, 2000; Zawojewski & Heckman, 1997; Zawojewski & Shaughnessy, 1999). In comparison, a limited number of studies have focused on teachers' understanding of average (Batanero, Godino & Navas, 1997; Cai & Gorowara, 2002; Callingham, 1997; Gfeller, Niess & Lederman, 1999; Groth & Bergner, 2006; Leavy & O'Loughlin, 2006).

Research concerning teachers' understanding of averages have focused entirely on the arithmetic mean (Batanero, Godino & Navas, 1997; Callingham, 1997; Gfeller, Niess & Lederman, 1999; Leavy & O'Loughlin, 2006); on the mean, median, and mode (Groth & Bergner, 2006); and on pedagogical knowledge of the concept of average (Cai & Gorowara, 2002). The two studies most closely related to this paper were conducted by Groth & Bergner (2006) and Leavy & O'Loughlin (2006). Both of these studies were conducted with preservice teachers (PSTs) whereas the research presented in this paper involves inservice teachers.

Groth & Bergner (2006) investigated 46 PSTs' understanding of the mean, median, and mode. In particular, the researchers focused on the PSTs' understanding of the differences and similarities between the three measures of center. The SOLO taxonomy (Biggs & Collis, 1982) was used to differentiate the 46 teachers into four categories: unistructural/concrete symbolic, multi-structural/concrete symbolic, relational/concrete symbolic, and extended abstract. Eight teachers were classified at the unistructural/concrete phase as they were only able to provide definitions of the various measures of center. Twenty-one teachers were at the multistructural/concrete symbolic phase. These teachers realized that the measures represented a mathematical object of importance rather than simply being the result of a procedure. Thirteen teachers were at the relational/concrete symbolic as they realized that measures of center tell us what is "typical" about a certain set of data. Finally, three teachers were at the extended abstract phase as they were able to articulate when one measure of center would be more appropriate or useful than another.

Although the study presented in this paper is related to the work of Groth & Bergner (2006), it is most aligned and extends the work of Leavy & O'Loughlin (2006). Leavy & O'Loughlin (2006) attempted to capture an in-depth understanding of PSTs' understanding of the mean. Their study was conducted with 263 undergraduate preservice elementary school teachers in Ireland. The researchers posed five tasks to all participants as well-conducted clinical interviews with 25 of the students. These interviews were meant to further examine the conceptual understanding of the PSTs. One group of individuals was chosen based on their responses to the five tasks whereas another group was chosen randomly (specific numbers in each group were not reported).

Although all five tasks were classified as assessing conceptual understanding of the mean, only one measured PST knowledge at the depth and type of understanding described in the GAISE framework. The first task focused on the recognition that finding the mean is an acceptable tool to use when comparing data sets with varying data points; however, it did not consider other measures of center as being appropriate. The second task involved the calculation of a weighted mean, which would be considered procedural. The third and fourth tasks would also be considered procedural in that they involved the participants constructing a data set to have a specific mean—which was provided to them. The final task involved participants estimating the mean based on visual characteristics of the distribution.

The first task of choosing the mean as a functional tool to compare data sets revealed some level of conceptual understanding by the participants. Although this task showed that several participants were able to find the mean to compare the data sets, it is possible that this type of question leads itself to a common misconception. This misconception involves the assumption that the only suitable measure of center or method for comparison is the mean as it did not allow or account for other measures of center.

The final task involved two parts. The first part had participants view a line plot and (1) indicate if a mean could be found based solely on that information; (2) if so, find the mean. The second part involved participants indicating what the mean represents in a distribution. In the first part of the task, only 3% of respondents utilized an approach that showed some level of conceptual understanding of the mean. Both parts of this task revealed that many participants confused the concept of the mean with the other measures of center (median and mode). It also revealed that participants generally (52%) view the term mean as synonymous with "average."

Leavy and O'Loughlin claimed that the majority of PSTs in their study did not possess much conceptual understanding of the mean; however, these claims were made while investigating tasks that mainly measured procedural understanding. Their study still adds value to the literature in regard to PSTs' understanding in relation to the mean. The results presented in this paper focus more on a conceptual understanding of the mean as well as the median as a method for exploring quantitative data.

In this paper, teachers' understanding is discussed within the context of procedural and conceptual understanding. "One kind of procedural knowledge is a familiarity with the individual symbols of the system and with the syntactic conventions for acceptable configurations of symbols. The second kind of procedural knowledge consists of rules or procedures for solving mathematical problems" (Hiebert & Lefevre, 1986, p. 8). On the other hand, "conceptual knowledge is characterized most clearly as knowledge that is rich in relationships. It can be thought of as a connected web of knowledge, a network in which the linking relationships are as prominent as the discrete pieces of information" (Hiebert & Lefevre, 1986, pp. 3–4).

In light of the suggestions offered by the GAISE framework for the increased level of sophistication at which statistical topics should be covered, this paper presents the results of a study that sheds light on the current status of elementary school teachers' understanding of the mean and median. An analysis of inservice teachers' current knowledge of these topics may influence the way future teachers are prepared to teach statistics at the elementary school level. This study attempted to answer the following research question:

What understanding of the mean and median do three exemplary elementary school teachers possess?

Method

A case study involving three elementary school teachers was conducted in a school district in the USA. The results presented in this paper are from a larger study that explored elementary school teachers' understanding in other areas of statistics as well as how their interaction with a particular curriculum and assessment instruments influenced their awareness of their knowledge (see Jacobbe, 2007).

Participants, Method of Selection, and Procedure

Three teachers were involved in the study—Ms. Alvin, Ms. Brown, and Ms. Clark. Ms. Brown and Ms. Clark taught grade 3 while Ms. Alvin taught grade 4. Ms. Alvin was the most senior of the three with 9 years experience. Ms. Brown and Ms. Clark had 5 and 4 years experience, respectively. Beyond an introductory statistics course, the three teachers in this study did not have any experience during their teacher preparation program that specifically prepared them to teach statistics.

Initial contact was made with the three teachers through a gatekeeper, the district supervisor for mathematics and science education. Several teachers were observed, interviewed, and surveyed before deciding on three teachers as the focus of this case study. The study was limited to three participants because of convenience sampling and the length of time necessary to visit, interview, and assess the teachers at the depth involved in this study. All three teachers were highly recommended by their principals and district supervisor as they were viewed as exemplary teachers of mathematics. Although generalizations cannot be made because of the small sample size involved in this study, the results provide a glimpse into what may be expected of exemplary elementary school teachers and complement the work of other researchers.

Over the course of 14 months, each of these teachers participated in interviews, completed questionnaires and assessments, and allowed the researcher to observe their classroom at least 12 times. The results reported in this paper are from teachers' responses to questions on some of the nine instruments involved in the case study.

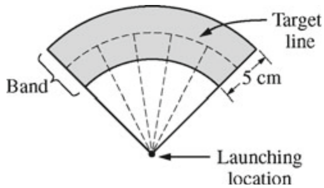
Instruments

Statistical Content Interview. The Statistical Content Interview included four questions related to the mean, four questions related to the median, one question concerning what the concept of an average, and one question relating the mean and median. Participants were also asked to provide a data set where the median would be a more appropriate measure of center than the mean. For the four questions related to the mean and median, participants were first asked to define the terms. If they were able to define them, they were asked the following follow-up questions: (1) How do you find the mean/median? (2) Is there more than one way? (3) What information does this value tell you?

Catapult Question. The catapult question was a question from the released 2006 AP® Statistics exam. This particular question was chosen because it pulled together how measures of central tendency and variation can be used to make a decision and was judged to be accessible by elementary school teachers. It was confirmed by an independent reviewer who was an assessment specialist in statistics that this question could be answered with an understanding at level A as described in the GAISE framework. The question was presented as displayed in Figure 1.

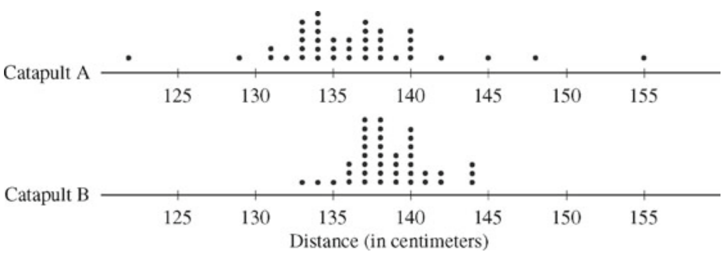
In order to ensure that the context was not confusing for the participants, the description was read and explained to them. After the entire context was worked through, participants were asked if they understood the situation and

Two parents have each built a toy catapult for use in a game at an elementary school fair. To play the game, students will attempt to launch Ping-Pong balls from the catapults so that the balls land within a 5-centimeter band. A target line will be drawn through the middle of the band, as shown in the figure below. All points on the target line are equidistant from the launching location.



If a ball lands within the shaded band, the student will win a prize.

The parents have constructed the two catapults according to slightly different plans. They want to test these catapults before building additional ones. Under identical conditions, the parents launch 40 Ping-Pong balls from each catapult and measure the distance that the ball travels before landing. Distances to the nearest centimeter are graphed in the dotplots below.



1. Explain the context in your own words.
2. What type of graphical displays are these?
3. What information can you find based on what is presented (i.e. mean, median, mode, range, and standard deviation)?
4. If the parents want to maximize the probability of having the Ping-Pong balls land within the band, which one of the two catapults, A or B, would be better to use than the other? Catapult A or B can be placed anywhere parents desire to maximize their chances of landing balls within the 5 cm band. Justify your choice.
5. Using the catapult that you chose in question 4, how many centimeters from the target line should this catapult be placed? Explain why you chose this distance.

Figure 1. Catapult questio```n

were then asked to explain it back to the researcher in the first question. This particular task is similar to Leavy & O'Loughlin's (2006) fifth task that asked participants to reason about a distribution presented in a line plot to determine where the mean may be located. This question differs in that it does not direct the participants to use the mean.

Card-Sorting Tasks. Participants were also asked to use data displays to analyze information about measures of central tendency. The card-sorting tasks involved three distributions—one normal distribution and two skewed distributions. In regard to measures of central tendency, the teachers were asked to examine the histograms and (1) indicate whether or not they could use the information to find the mean and the median and (2) arrange the measures of center for the distributions from least to greatest based on the value of the mean and median. The distributions were displayed as shown in Figures 2, 3, and 4.

Scales on the horizontal axis were intentionally not provided. This decision was made in an attempt to have participants focus on the shape of the distribution rather than on specific values. If participants seemed to struggle with the lack of values on the horizontal axis, a fictitious context was introduced. This context was the notion of test scores ranging from 0 to 100. Again, this task is similar to Leavy & O'Loughlin's (2006) fifth task that asked participants to reason about a distribution presented in a line plot to determine where the mean may be located. This particular series of tasks differs from their task in that it focuses on more than the mean and requires participants to make comparisons among various measures of center based solely on the shapes of the distributions.

Results and Discussion

During the interviews, all three teachers were able to describe an appropriate method for determining the mean and median. Teachers were also asked to describe the difference between the mean and median. In

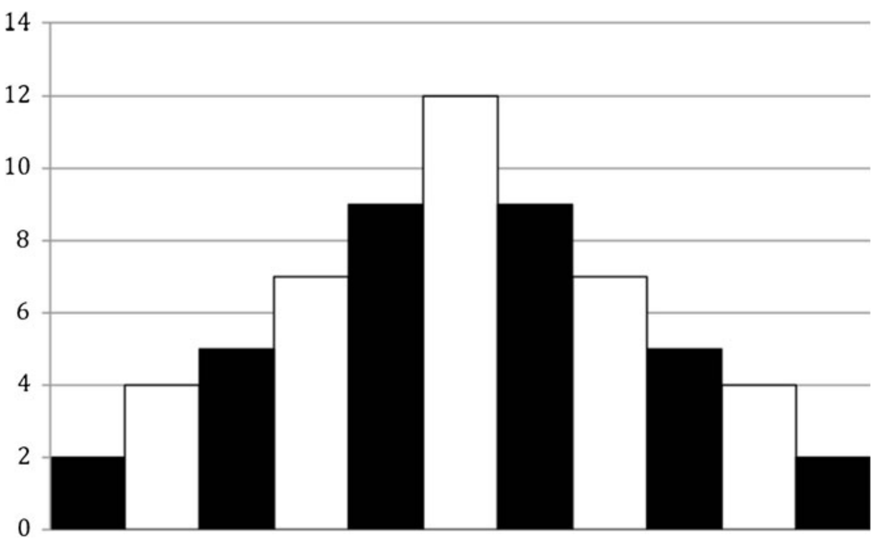


Figure 2. Distribution A

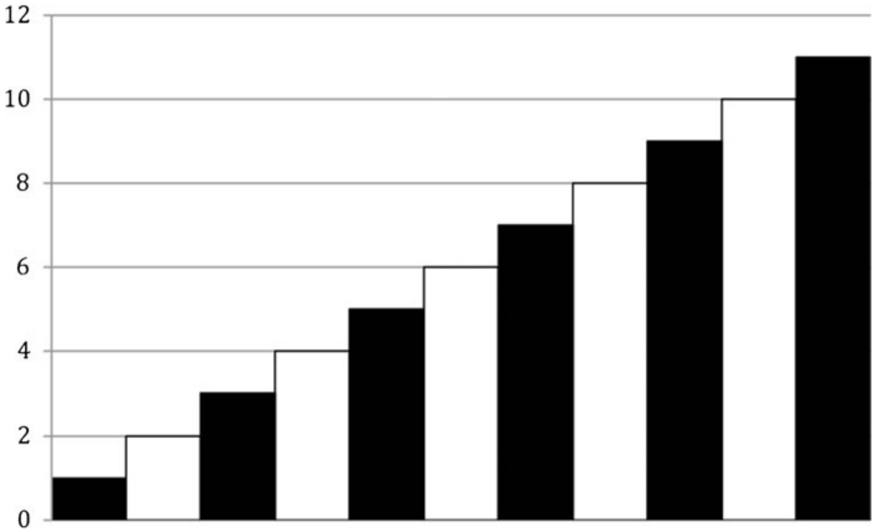


Figure 3. Distribution B

this line of questioning, two of the three teachers had difficulty explaining what these measures of center represent. Ms. Clark indicated how to calculate such measures rather than explaining what they represent. Her response is shown below and provides further evidence of how these teachers understood average as an algorithm.

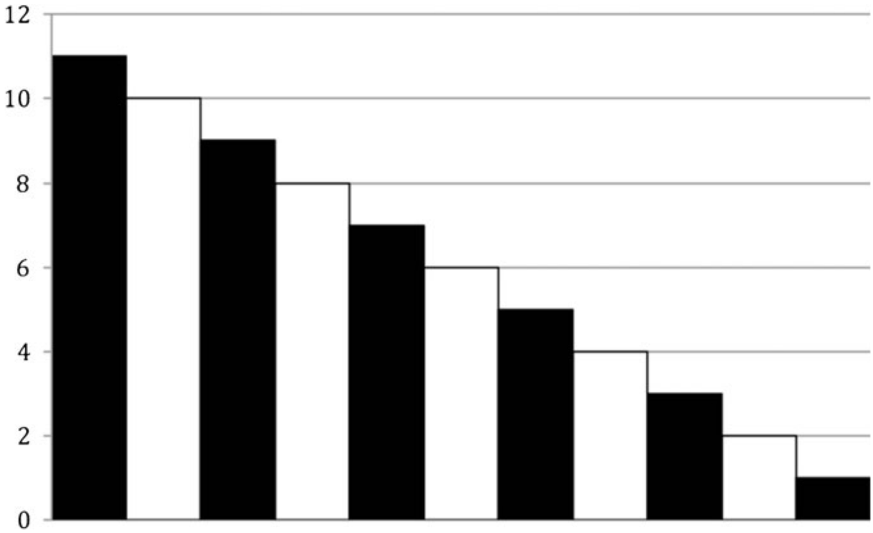


Figure 4. Distribution C

Researcher: What does the mean represent?
Ms. Clark: The mean is another word for average. So if you had the numbers, 1, 2, 3, 4, 5, then the mean would be 1+2+3+4+5 divided by 5.
Researcher: What does the median represent?
Ms. Clark: The median is the middle number, so if the numbers were 1 through 5 the median would be 3.

Ms. Brown was confused regarding the difference between the two measures of center. An example of her response in regard to the mean and median is shown below.

Researcher: What does the mean represent?
Ms. Brown: The average.
Researcher: What does the median represent?
Ms. Brown: The usual amount over a range of numbers.
Researcher: What does the median represent?
Ms. Brown: The number that is in the middle…but wait…they are the same. No, not really.
Researcher: So, what is the difference between the mean and the median?
Ms. Brown: I don't know the difference.

Based on the responses presented above, Ms. Brown and Ms. Clark appear to possess procedural knowledge of the mean and median; however, they did not possess conceptual knowledge of these topics.

The most conceptual response to these questions came from Ms. Alvin, as shown in the transcript comments below. The initial response indicates that Ms. Alvin could calculate the measures of center but could not provide more of an explanation. However, upon further probing, Ms. Alvin seemed to be working toward a conceptual description of the difference between the mean and median.

Researcher: What does the mean represent?
Ms. Alvin: I don't know.
Researcher: What does the median represent?
Ms. Alvin: I don't know… I don't know how to explain it. I guess I just know how to do it. I think that explaining it is difficult. It is easier just to show.
Researcher: What is the difference between the mean and the median?
Ms. Alvin: The median is finding the middle of all the data you have collected. You are not…I do not know…you have all the information there, but you do not manipulate the numbers to get one number. I don't know. All I can say is…the difference is when you are finding the average you are taking all of the numbers and manipulating them to get one number that represents the whole group and to find the median you still have that information, you're just finding the one that falls in the middle.

As discussed in the GAISE framework, Ms. Alvin's line of thinking is moving toward an understanding that the mean is influenced by every

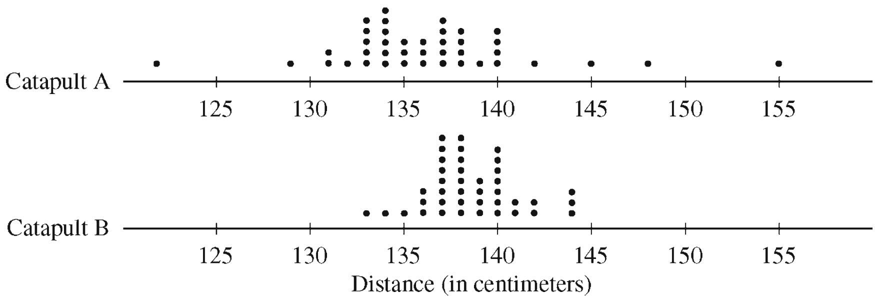


Figure 5. Dot plots from catapult task

point of data in the data set. In other words, the median is more resistant to outliers.

Another interesting response regarding the usefulness of the median was provided by Ms. Clark.

Researcher: What is the median useful for?
Ms. Clark: The median, being the middle number, would be useful…I don't really know. I guess just to know what a median is. I don't know why you would really need to know what the middle number is, but I guess to know how many times something is done or halfway.
Researcher: Could you give me an example of a set of data where the mean would be more useful than the median?
Ms. Clark: Since I really don't know what the purpose of the median is, the mean would be more important to me in any situation.

Ms. Clark's response again reveals that although she possessed a procedural knowledge of the median, she was unable to provide a reason why such a measure would be useful. Teachers' responses to the catapult question revealed other aspects of their understanding of the mean and median.

Performance on the Catapult Question

The catapult question involved the teachers examining comparative dot plots in order to choose the best catapult for landing ping-pong balls within a certain band of a target line for a game. Once they chose a catapult, they were to choose a location of where to place the catapult. The first task was based on an understanding of reducing variability, and the second was based on them understanding the concept of center. The context was thoroughly explained to the teachers, and they were asked to explain the context back to the researcher, prior to their choosing a catapult. The dot plots were displayed as in Figure 5.

Two of the three teachers (Ms. Alvin and Ms. Clark) were able to choose a catapult based on reducing the variability, and all three were all able to make a reasonable estimate of where to place the catapult. The teachers were able to make these decisions based on visualizing a balance point along the distribution. This type of task seems to involve a deeper understanding of center than a task where teachers are simply asked to perform a procedure. Ms. Clark's response shown below provides an example of this line of thinking. These findings differ from Leavy and O'Loughlin's results on their fifth task and indicate teachers that have difficulty with some aspects of procedural knowledge involving the mean or median may still be able to exhibit certain levels of conceptual knowledge.

Researcher: Using the catapult that you chose [before], how many centimeters from the target line should this catapult be placed? Explain why you chose this distance.
Ms. Clark: I would place it at either 136 or 137 centimeters from the target line. The majority of the ping pong balls… I guess I am looking for accuracy… so the majority of the ping pong balls would land… looking at catapult B most of the ping pong balls landed at 137 so give or take five centimeters… I think that placing it about 137 or 138 centimeters away from the target line would be appropriate.

As can be seen from Ms. Clark's thinking, she was continually shifting the center in her mind to try to "capture" as many ping-pong balls as possible. This line of thinking also shows that although teachers may lack some level of conceptual knowledge regarding the mean and median, they are able to visualize the concept and procedure as a balancing process, similar to what is introduced at level B of the GAISE framework.

Performance on Card-Sorting Task

Recall that the card-sorting tasks involved three distributions—one normal distribution and two skewed distributions (Figures 2, 3, and 4). The teachers were asked to examine the histograms and (1) indicate whether or not they could use the information to find the mean and the median and (2) arrange the measures of center for the distributions from least to greatest based on the value of the mean and median.

All three teachers indicated that the mean and median could be found using these distributions. However, only Ms. Clark was able to describe how to find the measures.

Researcher: Could you use the information in the displays to determine the mean?
Ms. Clark: I think I could do it; it might take a piece of paper. You had said they were on the same scale. If I knew these numbers I could calculate the mean by adding up all the

values and dividing by the total amount.
Researcher: Could you use the information in the displays to determine the median?
Ms. Clark: Yes, I would list all the numbers and find the middle number.

Notice that Ms. Clark always reverted back to the procedures she knew to calculate the mean and median. It should also be noted that, based on simply a histogram, one cannot determine the mean or median exactly. This is due to the fact that each bar of a histogram represents a range of values. For example, if one bar was presenting test scores in a histogram and scores ranged from 50 to 100, the first bar of the histogram may be used to display the frequency of scores that fell between 50 and 60. A histogram is used to display data and also give a sense of the overall shape of the distribution. Not one of the three teachers mentioned the shapes of the distributions in their responses. The teachers also were unable to relate the mean to the median (i.e. noting that the median would be equal, greater than, and less than the mean in distributions A, B, and C, respectively). Although this is a difficult question, none of the teachers had a strategy for exploring it. However, the teachers were able to use the shapes of the distributions to order the cards from least to greatest for the mean and median. Ms. Alvin was able to provide a description of how to organize the distributions from least to greatest for the means.

Researcher: Assuming these displays are on the same scale, place the cards from least to greatest according to their means. So place these cards based upon which distribution would have the smallest mean, which would have a mean in the middle of the other two, and which would have the greatest mean.

Ms. Alvin arranged the cards as follows: C, A, B

Researcher: Why did you order the cards this way?
Ms. Alvin: If these represent scores on a test from 50 to 100, then the 100 (pointing at Distribution C) would have the fewest amount. In Distribution B, there are more 100 s than any other score. Since Distribution A seems to have them equally spaced throughout, then I am leaving that one in the middle.

Ms. Alvin was able to correctly utilize the shapes of the distributions to order the cards from least to greatest according to the mean. Ms. Alvin and Ms. Clark were also able to successfully order the cards according to the median (i.e. distribution C, distribution A, distribution B). However, Ms. Brown was unable to order the cards in this manner.

Researcher: Can you order the distributions according to which distribution would have the smallest median, which would have the median in between the other two distributions,

and which would have the largest median?
Ms. Brown: I don t really know. I am looking at this right here (Pointing to the highest [middle] bar in Distribution A). The median is the middle number. So the median for Distribution A would be in this bar.
Researcher: What about for the other ones?
Ms. Brown: These on the ends (Pointing to the highest bars). Oh, oh, oh! No these in the middle (pointing to the middle bars in both distributions B and C). But they are the same. I got it. I am trying to rearrange them.
Researcher: How are you trying to rearrange them?
Ms. Brown: By putting the highest bar in the middle for B and C, like it is in A.
Ms. Brown: They would all be the same?
Researcher: All the medians would be the same? Where would they be?
Ms. Brown: Well, I am looking at it. The middle number is in the middle bar (the 6th bar from the left). So all the medians would be the same.

Ms. Brown did not acknowledge that the median of the distributions would be influenced by the way the data were distributed. In other words, distribution C would have the smallest median since there are more values on the lower end of the scale, distribution B would have the largest median since there are more values on the higher end of the scale, and the median of distribution A would be in between the medians of distributions C and B. Again, it should be noted that these values cannot be determined exactly, but it is possible to use the shapes of the distributions to determine in which bar the medians are located.

The results presented in this paper show that the teachers involved in this study lacked a connection between the procedures for finding the mean and median and what these measures of center actually represent within a particular context. However, at times the teachers showed some conceptual knowledge of the mean or median. Similar to the findings of Russell & Mokros (1991), these findings also speak to the importance of students (and teachers) possessing an understanding of average that is not dominated by an algorithm but focused on the underlying concepts those algorithms represent.

The results of this study add to the work of Leavy & O'Loughlin (2006) most notably in that it explored inservice teachers. Furthermore, this study involved an in-depth exploration of teachers' understanding of the mean and median. Four of the five tasks used in Leavy and O'Loughlin's study involved performing some type of computation surrounding the mean. This study involved tasks more similar to their fifth task and really focused on the conceptual understanding of teachers with respect to both the mean and median. In many ways, the results of this study support the conclusions of Leavy and O'Loughlin in that they identify weaknesses in the teachers' understanding and provide evidence that they do not possess much conceptual understanding. However, the results from the catapult and card-sorting task seem to contradict the results Leavy and O'Loughlin received on their fifth task, which involved teachers visualizing the concept of the mean. In this study, the teachers showed evidence of conceptual knowledge related to understanding distribution. In regard to the catapult task, this knowledge seemed to be supported by the task being situated within a context. Both studies have highlighted many of the gaps teachers may have in relation to the mean.

The three teachers involved in this study do not possess knowledge of the mean and median as outlined in the GAISE framework. It is important to note that this lack of knowledge is not due to the three teachers' inability to understand statistics but rather due to a lack of content exposure as described in the GAISE framework. If the sophisticated level of understanding described by the authors of the GAISE framework is to be realized by K–12 students, it is important that teachers are prepared to teach statistics at this level. Since these expectations are relatively new, most PSTs likely have not had sufficient experiences during their K–12 schooling to develop such an understanding.

Implications and Conclusion

The results of this study have implications for all countries concerned with statistics education. The situation in the USA is very similar to situations involving statistics education across the globe. The expectations for teaching statistics are being increased without first addressing teachers' preparation to teach the content.

As exhibited in this paper, it cannot be assumed that teachers understand the material at sufficient depth to teach statistical content effectively. Consequently, PSTs should be introduced to these types of activities during their preparation programs and inservice teachers should be provided with professional development opportunities. In both settings, preparation programs and professional development, the education community should pull from existing resources (e.g. Teach-STAT, 1996a, b) that have been used with teachers and evidence exists regarding their impact on teachers' content and pedagogical content knowledge (Friel & Bright, 1998). Activities that address the mean as a balance point should be specifically selected as these will help inform future teachers' of contexts where the median may be a more representative measure of center than the mean and vice versa.

Despite some of the setbacks, the study also revealed reasons for optimism. By the end of the study, all three teachers had acknowledged an awareness of their lack of content knowledge in the area of statistics and a desire to receive professional development focused on this particular content strand. From the larger study, it was clear that although the teachers did not experience lasting gains in their understanding of essential topics in statistics, it caused them to reconsider the suitability of their own content knowledge.

As Gal (2004) points out, one of the biggest obstacles toward the inclusion of statistics in the curriculum is teachers' (and society's in general) disposition toward the discipline of statistics. If teachers have only been exposed to statistical content as a discipline entrenched in procedures rather than the conceptual underpinnings of the processes, then they will not realize that there is a gap in their knowledge. For all the teachers in the case study, the interaction caused the teachers to question their dispositions toward statistics. In other words, this interaction problematized the awareness of their knowledge. The researcher became aware of this change through an interview. The participants in this study were prompted to indicate whether they would prefer professional development focused on (1) statistical activities they could do with students or (2) content they would need to understand in order to teach statistics. At the beginning of the study, all three teachers preferred activities they could use in their classrooms. At the end of the study, all three identified an interest in learning more content for teaching statistics effectively.

Once teachers recognize new viewpoints or what may be lacking in their own understanding, problematization occurs. Problematizing teachers' knowledge is essential for professional development to be successful in changing teachers' preparedness for teaching (Cobb & Bauersfeld, 1995). Teachers who realize they have a lack of understanding in a particular area are more likely to benefit from professional development focused on content. The teachers involved in this study provide an example that illustrates the importance of such a realization. With the problematization they experienced, the teachers would be more likely to absorb the content introduced during sustained professional development that is fundamental to true understanding and the processes of statistical inquiry.

REFERENCES

Batanero, C., Cobo, B., & Diaz, C. (2003). Assessing secondary school students' understanding of averages. *Proceedings of CERME 3*, Bellaria, Italia. Online: www.dm.unipi.it/~didattica/CERME3/.

Batanero, C., Godino, J. & Navas, F. (1997). Concepciones de maestros de primaria en formción sobre promedios (Primary school teachers' conceptions on averages). In H. Salmerón (Ed.), *Actas de las VII Jornadas LOGSE: Evaluación Educativa* (pp. 310–340). Granada, Spain: University of Granada.

Biggs, J. B. & Collis, K. F. (1982). *Evaluating the quality of learning: The SOLO taxonomy*. New York: Academic Press.

Cai, J., & Gorowara, C. C. (2002) Teachers' conceptions and constructions of pedagogical representations in teaching arithmetic average. In B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching Statistics*. Cape Town, South Africa: International Statistical Institute and International Association for Statistical Education. www.stat.auckland.ac.nz/~iase/publications.

Cai, J. & Moyer, J. (1995). Beyond the computational algorithm: Students' understanding of the arithmetic average concept. In L. Meira (Ed.), *Proceedings of the 19th Psychology of Mathematics Education Conference* (Vol. 3, pp. 144–151). Recife, Brazil: Universidade Federal de Pernambuco.

Callingham, R. (1997). Teachers' multimodal functioning in relation to the concept of average. *Mathematics Education Research Journal, 9*, 205–224.

Cobb, P. & Bauersfeld, H. (Eds.). (1995). *The emergence of mathematical meaning—interaction in classroom cultures*. Hillsdale, NJ: Lawrence Erlbaum.

Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M. & Scheaffer, R. (2007). *Guidelines for assessment and instruction in statistics education (GAISE) report*. Alexandria, VA: American Statistical Association.

Friel, S. N. & Bright, G. W. (1998). Teach-STAT: A model for professional development and data analysis for teachers K-6. In S. Lajoie (Ed.), *Reflections on statistics: Learning, teaching, and assessment in grades K–12* (pp. 89–117). Mahwah, NJ: Erlbaum.

Gal, I. (2004). Statistical literacy: Meanings, components, responsibilities. In D. Ben-Zvi & J. Garfield (Eds.), *The challenges of developing statistical literacy, reasoning, and thinking* (pp. 47–78). Dordrecht, The Netherlands: Kluwer.

Garcia, C., & Garret, A. (2006). On average and open-end questions. In A. Rossman & B. Chance (Eds.), *Proceedings of the Seventh International Conference on Teaching Statistics*. Salvador (Bahia), Brazil: International Association for Statistical Education. www.stat.auckland.ac.nz/~iase/publications.

Gfeller, M. K., Niess, M. L. & Lederman, N. G. (1999). Preservice teachers' use of multiple representations in solving arithmetic mean problems. *School Science and Mathematics, 99*, 250–257.

Groth, R. E. & Bergner, J. A. (2006). Preservice elementary teachers' conceptual and procedural knowledge of mean, median, and mode. *Mathematical Thinking and Learning, 8*, 37–63.

Hiebert, J. & Lefevre, P. (1986). Conceptual and procedural knowledge in mathematics: An introductory analysis. In J. Hiebert & P. Lefevre (Eds.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 1–27). Hillsdale, NJ: Lawrence Erlbaum.

Jacobbe, T. (2007). Elementary school teachers' understanding of essential topics in statistics and the influence of assessment instruments and a reform curriculum upon their understanding. Online: www.stat.auckland.ac.nz/~iase/publications.

Leavy, A. & O'Loughlin, N. (2006). Preservice teachers understanding of the mean: Moving beyond the arithmetic average. *Journal of Mathematics Teacher Education, 9*, 53–90.

McGatha, M., Cobb, P., & McClain, K. (1998). An analysis of students' statistical understandings. *Paper presented at the Annual Meeting of the American Educational Research Association*, San Diego, CA.

National Council of Teachers of Mathematics (1991). *Professional standards for teaching mathematics*. Reston, VA: NCTM.

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

Russell, S. J. & Mokros, J. R. (1991). What's typical? Children's ideas about average. In D. Vere-Jones (Ed.), *Proceedings of the Third International Conference on Teaching Statistics* (pp. 307–313). Voorburg, The Netherlands: International Statistical Institute.

Scheaffer, R. (1986). The quantitative literacy project. *Teaching Statistics, 8*(2), 34–38.

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. 465–493). New York: Macmillan.

Shaughnessy, J. M. (2007). Research on statistics learning and reasoning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 957–1009). Charlotte, NC: Information Age.

Teach-STAT (1996a). *Teach-STAT: Teaching statistics grades 1–6: A key for better mathematics. The University of North Carolina Mathematics and Science Education Network*. Palo Alto, CA: Dale Seymour.

Teach-STAT (1996b). *Teach-STAT for statistics educators. The University of North Carolina Mathematics and Science Education Network*. Palo Alto, CA: Dale Seymour.

Watson, J. M., Callingham, R. A., & Kelly, B. A. (2007). Students' appreciation of expectation and variation as a foundation for statistical understanding. *Mathematical Thinking and Learning, 9*, 83–130.

Watson, J. M. & Moritz, J. B. (1999). The developments of concepts of average. *Focus on Learning Problems in Mathematics, 21*, 15–39.

Watson, J. M. & Moritz, J. B. (2000). The longitudinal development of understanding of average. *Mathematical Thinking and Learning, 2*, 11–50.

Zawojewski, J. S. & Heckman, D. J. (1997). What do students know about data analysis, statistics, and probability? In P. A. Kenney & E. A. Silver (Eds.), *Results from the sixth mathematics assessment of the National Assessment of Educational Progress* (pp. 195–223). Reston, VA: National Council of Teachers of Mathematics.

Zawojewski, J.S., & Shaughnessy, J.M. (1999). Data and chance. In P.A. Kenney & E.A. Silver (Eds.), *Results from the seventh mathematics assessment of the National Assessment of Educational Progress* (pp. 235–268). Reston, VA: National Council of Teachers of Mathematics.

School of Teaching and Learning
University of Florida
PO Box 1170482403 Norman Hall, Gainesville, FL 32611, USA
E-mail: jacobbe@coe.ufl.edu