Prospective Teachers' Procedural and Conceptual Knowledge of Mean Absolute Deviation

Randall E. Groth
Salisbury University
regroth@salisbury.edu

Abstract

The recommendation to study statistical variation has become prevalent in recent curriculum documents. At the same time, research on teachers' knowledge of variation is in its beginning stages. This study investigated prospective teachers' knowledge in regard to a specific measure of statistical variation that is new to many curriculum documents: the mean absolute deviation (MAD). Seventy-six prospective teachers participated in the study. Participants exhibited various procedural and conceptual characteristics in their thinking about the MAD. The majority were able to successfully select and carry out a procedure for computing the MAD. However, some had difficulty dealing with procedures for absolute deviations, and others confused the procedure for the MAD with the procedure for a different descriptive statistic. Conceptually, participants offered a variety of interpretations of the MAD, with some demonstrating deep understanding of the measure and others demonstrating shallower understanding or misconceptions. Those who demonstrated the strongest conceptual knowledge of the MAD also exhibited sound procedural understanding, suggesting that the two types of knowledge are intertwined in the process of fully understanding the measure.

Statistical variation has become widely acknowledged as an important object of study for school curricula because of the role it plays in statistical thinking. Cobb and Moore (1997) argued that the discipline of statistics exists because of the "omnipresence of variability" as seen in situations such as variation among individuals and repeated measurements. Wild and

Pfannkuch (1999) expanded on the notion of the omnipresence of variability, stating, "Variation is an observable reality. It is present everywhere and in everything. Variability affects all aspects of life and everything we observe. No two manufactured items are identical, no two organisms are identical or react in identical ways" (p. 235). Garfield and Ben-Zvi (2008) argued that the study of variability should permeate statistics curricula, and that understanding variability is necessary for learning the concepts of distribution and statistical inference. Such observations about variability in relation to statistical thinking are reflected in its central role in various curriculum documents (Aliagea et al., 2005; Franklin et al., 2007; National Governors Association for Best Practices & Council of Chief State School Officers, 2010).

An important part of the study of statistical variation is its quantification. Statisticians have developed several measures to quantify the variability of data, such as the range, interquartile range, and standard deviation. In this report, I focus primarily on one particular measure: the mean absolute deviation (MAD). It is the mean distance between the individual data values in a data set and the mean of the data set. Recommendations to include the MAD as an object of study in middle school curricula are relatively new (Franklin et al., 2007, National Governors Association for Best Practices & Council of Chief State School Officers, 2010). A primary rationale for its inclusion in middle school curricula is that the MAD can serve as a precursor to the study of standard deviation at the secondary or college level (Kader & Mamer, 2008). The mathematical structure of the MAD is similar to that of standard deviation, but does not require calculations of squares or square roots. Because of this, the MAD might provide a more intuitive introduction to the idea of measuring the typical distance data values are from the center of a distribution.

The recent inclusion of the MAD in middle school curricula raises questions about teachers' knowledge. Sanchez, da Silva, and Coutinho (2011) observed that research on teachers' understanding of variation is scarce. They identified an urgent need for further research in this area. The need to understand teachers' knowledge of the MAD is particularly pressing at this time, as many sixth and seventh-grade teachers in the U.S. are now required to teach it as part of the *Common Core State Standards* (National Governors Association for Best Practices & Council of Chief State School Officers, 2010). In order to support teachers in implementing instruction related to the MAD, it is optimal for teacher educators and researchers to be able to anticipate reasoning patterns and potential difficulties teachers may have with the idea. The current scarcity of relevant research makes it difficult for them to do so. This study thus aims to fill a gap in the literature by helping to characterize teachers' knowledge of the MAD.

Procedural and Conceptual Knowledge of
Mean Absolute Deviation

One risk of introducing a new idea such as the MAD to school curricula is that teachers may opt to portray it in strictly procedural terms. For example, in the U.S., mathematics instruction is often best characterized as the process of repeating and applying procedures of low complexity, demonstrated by the teacher (Jacobs et al., 2006). Bakker (2004) observed a similar phenomenon in statistics classes, noting that students seemed to have been drilled to calculate measures such as mean, median, and mode and produce displays such as bar graphs without understanding how these tools could be profitably used in statistical analysis. He argued that well-rounded statistical understanding consists not only of procedural knowledge, but that conceptual knowledge is needed as well. Similarly, it has been argued that teachers need both conceptual and procedural knowledge if they are to help their students learn with understanding (Kilpatrick, Swafford, & Findell, 2001; Ma, 1999).

Defining Conceptual and Procedural Knowledge

Systematic study of teachers' conceptual and procedural knowledge requires working definitions for the two knowledge types. Hiebert and Lefevre (1986) offered the following influential definitions:

> 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. Relationships pervade the individual facts and propositions so that all pieces of information are linked to some network (p. 3).

> Procedural knowledge...is made up of two distinct parts. One part is composed of the formal language, or symbol representation system, of mathematics. The other part consists of the algorithms, or rules, for completing mathematical tasks (p. 5).

These definitions helped provide common language for discussing the two types of knowledge as they came under investigation by researchers.

Star (2005) suggested revisions to the Hiebert and Lefevre (1986) definitions. Star took issue with using the descriptor "rich in relationships" only for conceptual knowledge. He argued that procedural knowledge in some instances is also rich in relationships and connections, citing examples of procedural thinking that call for comprehension, flexibility, and critical judgment. Star believed that using the Hiebert and Lefevre definitions encouraged researchers to view procedural knowledge as a lesser degree of knowledge rather than a distinct knowledge type. By not restricting the "rich in relationships" descriptor to conceptual knowledge. Star hoped to

encourage more research on the nature of procedural knowledge. Star believed that the issue of knowledge type and knowledge quality needed to be disentangled to encourage more research focusing on the nature of procedural knowledge and its development.

Baroody, Feil, and Johnson (2007) formulated a response to Star's (2005) reconceptualization of procedural knowledge. They agreed with Star that both procedural and conceptual knowledge were important objects of study, and that procedural knowledge should not automatically be equated with shallow rote knowledge. They disagreed, however, on the nature of deep procedural knowledge. Baroody et al. argued that deep procedural knowledge comes about through connections to conceptual knowledge. Star (2007) responded that framing deep procedural knowledge in terms of conceptual knowledge once again encourages the field to view procedural knowledge as a lesser form of knowledge not worthy of study in its own right, stating, "procedural knowledge is valuable in and of itself, not solely because of its connections with and integration to conceptual knowledge" (p. 134). Star believed that carving out a distinct niche for procedural knowledge rather than framing it solely in terms of a secondary role to conceptual knowledge would encourage research foregrounding procedural knowledge.

The present study draws upon points of agreement between Star (2005) and Baroody et al. (2007) in the notions of conceptual and procedural knowledge it employs. Specifically, it attempts to avoid entanglement of the issues of knowledge quality and type by allowing for the possibility that procedural runs on a continuum from shallow to deep, and that the same holds for conceptual knowledge. Both types of knowledge are considered important because of their necessity to well-developed teacher content knowledge (Kilpatrick et al., 2001; Ma, 1999). The question of whether deep procedural knowledge must be grounded in conceptual knowledge, a point of disagreement between Star and Baroody et al., is left to be addressed by empirical data gathered during the study. As Baroody et al. noted, their disagreement with Star's view is best addressed empirically. Although the present study does not aim to completely settle the disagreement, it can continue toward its resolution within the specific context of prospective teachers' conceptual and procedural knowledge of the MAD.

Procedural and Conceptual Knowledge for the MAD: An Example

Given the foregoing discussion of conceptual and procedural knowledge, it is worthwhile to consider what each type of knowledge may look like within the context of the MAD. Consider the following set of scores from a test: 75, 82, 86, 90, 97. A procedure for computing the MAD would be as follows:

1. Compute the mean of the distribution: $\frac{75+82+86+90+97}{5} = 86$

2. Determine the absolute values of the deviations from the mean:
   |75 - 86| = 11; |82 - 86| = 4; |86 - 86| = 0; |90 - 86| = 4; |97 - 86| = 11
3. Compute the sum of the absolute deviations (SAD):
   11 + 4 + 0 + 4 + 11 = 30
4. Divide the SAD by the number of values in the distribution: 30 ÷ 5 = 6.

Jacobs et al. (2006) categorized problems as "high complexity" if they contained two or more sub-problems and required more than four decisions by students. They found high complexity problems to be rare in the U.S. mathematics lessons they studied, "accounting for averages of 7% of the problems per lesson in 1995 and 6% in 1999" (p. 17). According to the Jacobs et al. definition, the MAD procedure is of high complexity because it contains of two sub-problems (computing the mean and computing the SAD) and requires several steps to complete. Hence, learning the procedure for the MAD is likely to pose a substantial challenge for students because of its unusually complex nature.

As with many other procedures, it is conceivable that students could master the procedure for computing the MAD without understanding the significance of the measure. Mokros and Russell (1995), for example, demonstrated that students may know the algorithm for computing the arithmetic mean without understanding that the mean provides a measure of the center or typical value in a data set. An analogous situation could occur for the MAD: students might master the algorithm without understanding that the MAD is a measure of the typical distance data values in a given distribution lie from the mean of the distribution. Kader and Mamer (2008) described another potential aspect of conceptual knowledge of the MAD. They noted the importance of realizing that the SAD can be sufficient for comparing variability in two data distributions that have the same number of values, but that it often breaks down when data sets have different numbers of values. The SAD might suggest more variability in a given data set simply because it has more values, but the MAD accounts for differences in sizes of data sets by dividing by the number of values in the sets.

Method

Participants

The study participants were enrolled in my undergraduate statistics course for prospective PreK-8 teachers (Groth, 2012, 2013a). The course was designed to build participants' statistical knowledge for teaching (Groth, 2013b), which encompasses both subject matter knowledge and pedagogical content knowledge (Hill, Ball, & Schilling, 2008). The MAD was an important object of study during the course because it appears in the

Common Core State Standards for grades 6 and 7. This meant that several participants would be likely to teach the MAD directly to children after completing their teacher preparation programs. I believed that even those who would not directly teach it would benefit from studying it by gaining what Hill et al. (2008) called "horizon knowledge," that is, knowledge of what they were preparing children to study in later years. Data for this study were gathered during four different semesters that I taught the course. The total number of participants was 76, consisting of 67 females and 9 males. Participants' consent was obtained to use the work samples included in this report, and pseudonyms are used in place of real names in connection with the work samples.

Procedures

Each semester of the course included instruction on the procedural and conceptual characteristics of the MAD. To introduce the idea of statistical spread or variability, I asked participants to use their own strategies to describe the spreads of distributions They were then asked to compare their strategies for measuring spread with formal measures, including range, interquartile range, MAD, and standard deviation. An activity described by Garfield and Ben-Zvi (2008) was then used to analyze characteristics of the MAD in detail. In the activity, participants used horizontal bars to represent deviations from the means of data sets. The MAD was then portrayed as the average bar length. After completing the in-class activities, participants were assigned Kader and Mamer's (2008) teacher-oriented article on describing center and variability in distributions. The article included a discussion of computing the MAD and an explanation of why the MAD should be used rather than the SAD when comparing different-size groups.

Instrument

After reading the Kader and Mamer (2008) article, participants responded to the following task:

> Construct two different sets of data that have the same mean. The data sets should have different numbers of values. Compute the SAD and MAD for each set of data. Show your work. Explain what the SAD and MAD tell you about the sets of data.

The task was part of a set of writing prompts designed to help participants engage with the pedagogical content knowledge and subject matter knowledge elements of assigned readings from teacher-oriented journal articles (see Groth, 2012 for the full set of prompts). This particular prompt was designed primarily to elicit participants' subject matter knowledge by asking them to demonstrate what they had learned in class and by reading the assigned article. The prompt allowed for demonstrating of procedural knowledge by asking for computations of SAD and MAD, and it allowed

for demonstration of conceptual knowledge by asking for interpretations of the two measures within the content of a particular data set.

Data Analysis

The great degree of qualitative participant-to-participant variation observed as the MAD task was completed provided the impetus to compile the data from each semester of the course and analyze responses carefully against one another. To being the formal qualitative analysis, I loaded responses for all participants into Atlas.ti (Muhr & Friese, 2004). Initially, open coding (Corbin & Strauss, 2008) was used to allow codes enough flexibility and malleability to succinctly capture the essence of the knowledge displayed in the responses. Each response was tagged with descriptive qualitative codes to capture evident procedural and conceptual characteristics. Once a set of codes had been developed to describe the procedural and conceptual characteristics of the data, codes with similar characteristics were clustered into conceptual categories (Miles & Huberman, 1994). Initially, codes were grouped into the two board categories of "conceptual understanding" and "procedural understanding." As code clusters were refined, it was possible to identify additional clusters within the two broad categories. The qualitative analysis thus resulted in a cross-case analysis that identified salient conceptual and procedural thinking patterns in participants' responses.

Results

Characteristics of participants' knowledge of the MAD are sorted into two categories below: procedural and conceptual. This organizational scheme is not meant to imply that the characteristics reported in each category are strictly procedural or strictly conceptual. It is often difficult to neatly separate the two types of knowledge. Instead, the organizational scheme used below sorts the characteristics into categories based on which of the two types of knowledge they most strongly exemplify.

Procedural Knowledge

Sixty-four of the 76 participants chose a correct procedure for calculating the MAD. Of these 64 participants, 11 made minor computational errors when carrying out the procedure (e.g., subtracting or dividing incorrectly). One correct procedure was to calculate the SAD and then divide it by the number of values in the data set, as illustrated in the Kader and Mamer (2008) article that had been assigned. Jenny's response, show in Figure 1, is illustrative.

set 1 → 6, 4, 7, 7, 2, 4

mean: 6 + 4 + 7 + 7 + 2 + 4 = 30 ÷ 6 = ⑤

SAD: 6 to 5 = 1, 4 to 5 = 1, 7 to 5 = 2, 7 to 5 = 2, 2 to 5 = 3,
5 to 5 = 0
1 + 1 + 2 + 2 + 3 + 0 = ⑨

MAD: SAD ÷ number of values
9 ÷ 6 = ⑴.5⑴

set 2 → 3, 5, 5, 5, 7

mean: 3 + 5 + 5 + 5 + 7 = 25 ÷ 5 = ⑤

SAD: 3 to 5 = 2, 5 to 5 = 0, 5 to 5 = 0, 5 to 5 = 0, 7 to 5 = 2
2 + 0 + 0 + 0 + 2 = ④

MAD: SAD ÷ number of values
4 ÷ 5 = ⑴.8⑴

SAD: This tells you the distance of the data points from
the mean. This then tells you how much variation there
is in the data. In this case the SAD is greater in data
set 1.

MAD: This tells the average variation from the mean.
In this case the MAD is greater in data set 1.

Figure 1.
Jenny's response to the MAD task.

Another correct procedure involved identifying the value of the SAD *apart from* calculating the MAD, as illustrated in Katie's work in Figure 2.

hockey goals in 7 games

Mean: 3
MAD: 2
2
2
.
1
3
'2   √12 = .7
MAD = 1.7
SAD = 12

Team A

mean

Team B

Mean = 3
MAD = 0
2
2
4
.
1
2
2
'2   √12 = 1.7
MAD = 1.7
SAD = 12

The MAD is the typical deviation
away from the mean which shows
that both graphs the points differ
.7 away from the mean. The SAD
is the sum of the absolute deviations
from the mean and in both data
sets it is 12 which means as a whole
of the data set the data points are 12
away from the mean

Figure 2.
Katie's response to the MAD task.

Although the type of response shown in Figure 2 implicitly includes a calculation of the SAD as a step in determining the MAD, it explicitly attends first to calculating the average of the absolute deviations. The value of the SAD is provided as somewhat of afterthought, likely included to satisfy the requirements of the stated task rather than to show an important part of the MAD procedure.

Participants selecting incorrect algorithms exhibited a wide spectrum of thinking. Seven participants stated that the SAD is obtained by summing all of the values in a data set. Angie, for example, wrote the two data sets {1, 9, 3, 7} and {3, 12, 2, 6, 3, 1, 10, 3} along with the formula "SAD = The total sum of the given data." She then applied her formula to the two data sets to obtain a SAD of 20 for the first and 40 for the second. A related procedural error was to calculate the MAD by calculating the mean of the data. All five of the participants making this error had also determined the SAD by summing the values in their data sets. The two participants who computed the SAD by summing the data values but then computed the MAD correctly were among those who computed the SAD apparently just to fulfill the requirements of the task rather than identifying it as a step in computing the MAD. The first portion of Candace's response was illustrative:

> The two sets of data that we are going to look and it and solve are: 3, 10, 17 and -5, 15, 20...The SAD for the first set of data is thirty because when you add up all the numbers, the value is thirty. To calculate the MAD for the first set of data is you find the mean, and then subtract the mean number from your set of data values, finally you add up those numbers and divide by 3, which gives you your MAD. The MAD for your first set of data is 4.3.

Unlike most of the responses that over-simplified the SAD procedure to summing data values, Candace's response includes attention to deviation from the mean in computing the MAD.

There were three participants, like Candace, who made procedural errors in computing the SAD and MAD while still showing evidence of taking deviation from the mean into account. Sally wrote that the SAD and MAD were sets of deviations, and did not condense the deviations to a single value. For one of her data sets, she wrote {1, 3, 3, 3, 5}, calculated the mean of the data to be three, and then calculated the set of absolute deviations from the mean to be {2, 0, 0, 0, 2}. She identified the set of absolute deviations as both the MAD and the SAD of the data set. Betty's work was similar to Sally's in that she identified the set of absolute deviations to be the SAD, but different from Sally's in that Betty chose a correct procedure for computing the MAD. Edna was similar to Sally in that she identified the MAD as the set of absolute deviations, but different in that she calculated the SAD to be the number of points in the original data set. Although Sally, Betty, and Edna recognized the need to compute absolute deviations from the mean,

they did not consistently coordinate those computations with an algorithm for summarizing the absolute deviations with a single value.

Some of the participants who did summarize typical absolute deviation with a single value exhibited other types of procedural errors. Annie used deviations rather than absolute deviations in computing the MAD. As a result, when she summed the deviations, she obtained zero and concluded that the MAD was zero as well. Ryan computed the SAD by counting the number of values in the data set that did not have the same value as the mean. Nonetheless, he computed the MAD correctly because he did not think of the SAD as an intermediate step in its algorithm. The importance of deviation from the mean was acknowledged in Annie's and Ryan's responses even through inconsistencies with the conventional algorithms for the SAD and/or MAD were left unresolved.

Others whose algorithms for the SAD and MAD were incorrect appeared to confuse the procedures for computing SAD and MAD with algorithms for other measures of variability. Alice subtracted the low value from the high value in each of her data sets. Doing so gave her a value of 1 for the first data set and a value of 2 for the second. She then concluded that the SAD and MAD were greater for the second data set without using a correct algorithm for either one of them. Stephanie also focused on the high and low values in her data sets, but added them together rather than subtracting the smallest from the largest. She believed that adding these values produced the SAD, and she divided the SAD she computed by the number of values in each data set to compute the MAD. It appears that both Alice and Stephanie attempted to apply the algorithm for a different measure of variability, the range, to the computation of the SAD and MAD. The range was not the only measure of variability whose algorithm was misapplied. In two cases, participants used the algorithm for the MAD to compute the SAD. In another case, one participant, Laura, appeared to be trying to use the algorithm for standard deviation to compute the SAD, since she squared each deviation from the mean, added the squares together, and then computed the square root of the result. The fact that range and standard deviation were also introduced as measures of variability in class may have contributed to the selection of the incorrect algorithms for SAD and MAD in such cases.

Conceptual Knowledge

One of the main conceptual learning goals for the course in which the participants were enrolled was to understand that the MAD is unique among other measures of variability because it provides a measure of typical absolute deviation from the mean. When asked to explain what the MAD tells about the data, 31 of the 76 participants described it in terms of a measure of typical or average distance from the mean. All 31 of these participants were also among those who had employed a correct algorithm for computing

the MAD. Jenny's and Katie's explanations for the meaning of the MAD, shown in Figures 1 and 2, are illustrative. Jenny spoke of the MAD in terms of average variation, and Katie described it as a typical deviation after completing the appropriate calculations to determine the absolute deviations.

Several participants who spoke of the MAD as a measure of variation gave less thorough explanations of what it measures. Five stated that it measures variation from the mean without specifying that it measures *typical* variation. Kelsey, for example, gave the following analysis after calculating the MAD for the two data sets she had invented:

> The MAD of the data set tells us that data set 1's data points deviate further from the mean than data set 2's data points. The data is data set 1 is more spread out than data set 2's data.

Another ten participants spoke of the MAD only as a measure of variation without mentioning variation *from the mean* at all. Deborah, for example, stated, "The SADs and MADs show you how much variability there is in a data set. The higher number the SAD and MAD are the more variability there is." One of the participants who did not attend to variation from the mean, Patricia, also used categorical data for one of her invented data sets and attempted to compute its MAD and SAD by using category frequencies. Patricia's other data set, however, consisted of quantitative data for which she correctly calculated the MAD. All but one of the 15 respondents who spoke of the MAD as a measure of variation in these less precise ways used a correct algorithm to determine the MAD. The only one of the 15 who did not select a correct algorithm at some point was Edna, who believed the MAD to be a set of absolute deviations rather than a single value. Although such responses were correct in characterizing the MAD as a measure of variation, they did not have the same conceptual depth as those specifically identifying the MAD as a measure of typical distance from the mean.

Another of the conceptual learning goals for the course was to understand that the MAD is more useful than the SAD when comparing groups that are not the same size. Only six of the participants mentioned this conceptual characteristic of the MAD in responding to the task. Jenna was among these six students, and her work is shown in Figure 3. After computing the SAD for each data set she invented, she concluded that the set with more values had more variation. However, after computing the MAD, she realized that the MAD accounts for different-sized groups and changed her explanation accordingly, stating at the end of her response, "Obviously the set with more values will have the greater sum of differences. The MAD makes up for this difference." Part of the reason for the lack of attention to the MAD's usefulness for comparing unequal groups appears to be that when 29 of the participants invented their data sets for the task, they made the data sets the same size, even through the task asked for data sets of different sizes. Making the data sets have the same number of values took away an opportunity

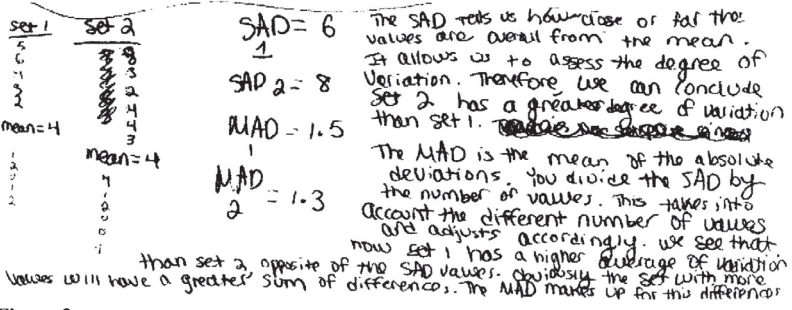to reflect on the relative usefulness of the SAD and MAD for comparing different-size groups.

Figure 3.
Jenna's response to the MAD task.

Not all participants who invented two data sets of different sizes, however, had strong conceptual explanations of the relationship between the MAD and SAD. Four participants expressed the belief that the data set with the larger SAD automatically has more variation. An example of this is shown in Tim's work in Figure 4.
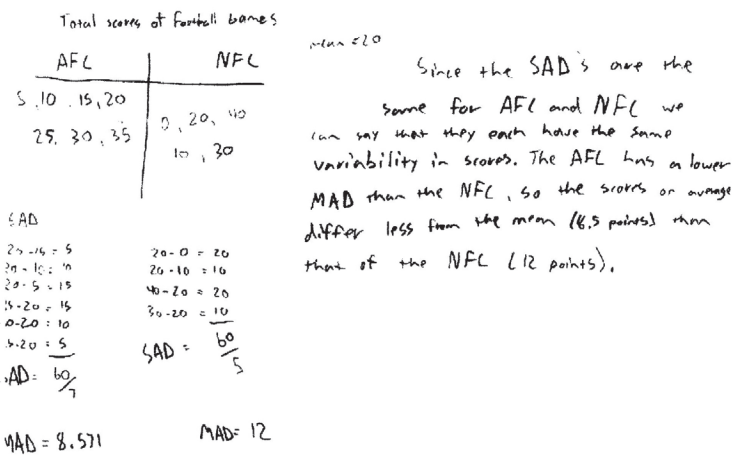

Figure 4.
Tim's response to the MAD task.

After computing the SADs for his two invented data sets to be the same, Tim concluded they had the same amount of variability because the SADs were the same. Then, after computing a lower MAD for the data set with more data values, he stated that the data values in that set differed less from the mean on average. This seemingly created a conflict between his interpretations of the meaning of the SAD and MAD, but he left the conflict un-

resolved. Tim's response stands in contrast to that of Jenna (Figure 3), who ultimately recognized that the MAD is better suited to compare variability when dealing with data sets of different sizes.

Some additional aspects of the relationship between the MAD and SAD proved to be difficult for participants to construct. One participant, Diane, computed the MAD using the algorithm for the median, computed the SAD with the algorithm for the MAD, and then wrote, "The MAD tells the median and the SAD is the average of how far away each point is from the median." Two participants said that if the SAD is larger in a data set, the MAD will be larger as well. Both of these participants created data sets with the same numbers of values and computed the MAD and SAD correctly, so this perceived relationship between the MAD and SAD held for the examples they had invented. The did not think beyond their examples, however, to test the viability of the generalization they had made. Another seven participants made no conceptual distinction between the MAD and SAD when interpreting them. One of these participants was Alice, who used an incorrect algorithm for the MAD. After computing the SAD by subtracting the lowest value in each set from the highest, and then dividing the difference by two to compute the MAD, she concluded, "The MAD tells me that data set B is more spread out from the mean, and therefore more spread in general. The SAD shows me the same, data set B is more spread." Some of those making no distinction between MAD and SAD had used correct algorithms to compute them. Marcy, for example, correctly computed the measures but then made no distinction in her interpretations of them, stating, "The SAD and MAD are ways to assess the degree of variation in data from the mean."

The meanings of MAD and SAD were also occasionally confused with meanings of other descriptive statistics. Both Diane and Annie interpreted the MAD to have the same properties as the median. Both also used incorrect algorithms for the MAD. As mentioned earlier, Diane simply used the algorithm for the median to compute the MAD. Annie used deviations rather than absolute deviations to compute it, and hence computed a MAD of zero. To interpret the zero value for the MAD, Annie stated, "Since the mean average deviation [*sic*] is 0, it means that about half of the numbers in the data set were above the mean and about half were below it." Another participant, Nancy, computed the MAD and SAD correctly but then made an incorrect generalization from the results of her computation in interpreting the measures. She stated, "Although the data sets are very different, the average deviations from the means are very similar. The same mean in the data set allowed for the MADs to be similar although the SADs themselves were twice the difference." Nancy's interpretation suggested that data sets with similar means tend to have similar MADs as well, rather than focusing on the MAD as a measure of average distance between the points in the data set and its mean.

Discussion

Figure 5 is a diagram summarizing common pathways participants took in solving the MAD task. Each node on the left side of the diagram represents a procedural knowledge characteristic, and each node on the right represents a conceptual knowledge characteristic. Procedural knowledge nodes show both correct and incorrect algorithms employed in responding to the first portion of the task. Conceptual knowledge nodes show degrees of conceptual understanding as well as misconceptions exhibited when participants were asked to interpret the meaning of MAD and SAD in the second portion of the task. The three conceptual knowledge nodes at the top of the right column suggest the possibility of deep conceptual knowledge of MAD (i.e., interpreting it as an indicator of typical distance from the mean) as well as shallower conceptual knowledge (i.e., describing the MAD as
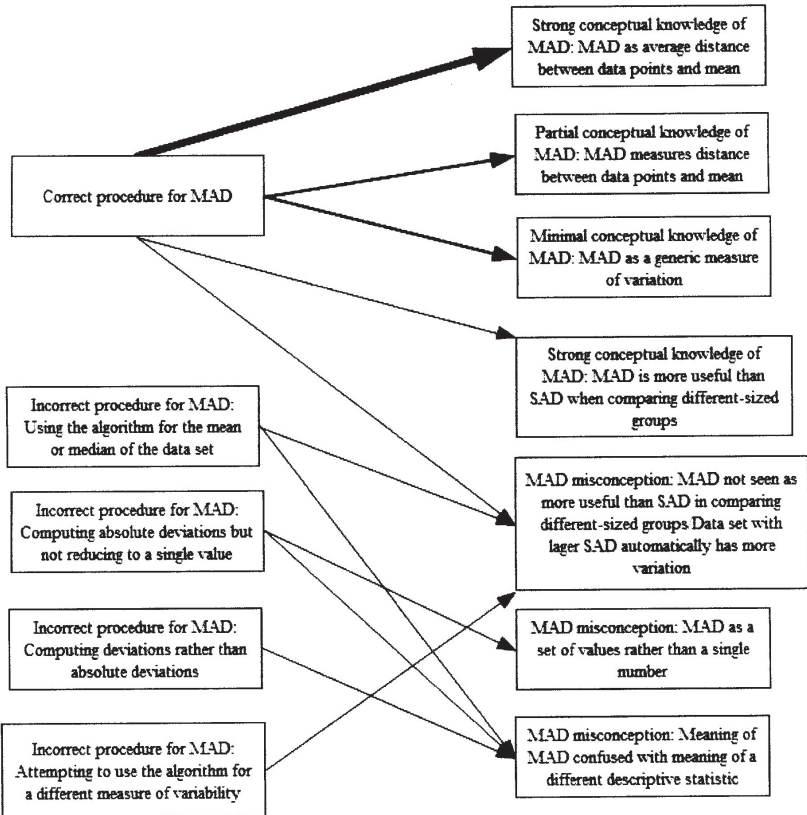
Figure 5.
Common pathways taken in solving the MAD task.

a more generic measure of variation). The arrows between the procedural and conceptual knowledge nodes represent pathways taken by participants in responding to the task. The thickness of each arrow roughly represents the relative frequency with which the indicated pathway was taken by participants

One salient feature of Figure 5 is that the most heavily traveled path taken in responding to the task was to use a correct algorithm for the MAD and then to exhibit strong conceptual understanding in interpreting it (31 of the 76 participants). Furthermore, the only participants exhibiting strong conceptual knowledge of the MAD were those who had first used a correct algorithm to compute it. Fourteen others who had selected correct algorithms exhibited lesser degrees of conceptual understanding, but at least did not show signs of outright misconceptions about the meaning of the MAD (although some did not identify the MAD as being more useful than SAD when comparing different-size groups). Those who did select incorrect algorithms were generally those who misinterpreted the meaning of the statistic (the one exception to this pattern was Edna's response, as described earlier). These data add a dimension to the discussion about the relationship between conceptual knowledge and procedural knowledge taken up by Star (2005) and Baroody et al. (2007). Star and Baroody et al. debated the question of whether deep procedural knowledge must be grounded in conceptual knowledge. The data from the present study suggest a different, but related question: In the case of the MAD, is strong *conceptual* knowledge necessarily grounded in strong *procedural* knowledge? If so, this provides strong support for Star's contention that more research foregrounding procedural knowledge is necessary.

To address the question of the relationship between conceptual and procedural knowledge in the case of the MAD, it can be helpful to re-examine the algorithm for the MAD in addition to the empirical data that have been generated. The MAD algorithm draws attention to the computation of the mean of the distribution as well as the distances of data points from the mean. These computations may provide significant cues about the meaning of the MAD, as past research has shown that it is often not natural for students to think of statistical variation in terms of distance from the mean (delMas & Liu, 2005). At the same time, a degree of conceptual understanding is required in thinking of the mean as a measure of center and thinking of subtraction and absolute value as measuring distances. The MAD algorithm then directs students to compute the arithmetic mean of the distances from the mean of the distribution, prompting students toward understanding the MAD as a one-number summary of variation in the distribution. Conceptual knowledge again plays a supporting role in this step of interpretation, as students must understand the mean as a measure of typical value to appreciate the MAD as the typical distance from the mean. Hence, it seems

that conceptual knowledge and procedural knowledge are intertwined in understanding the MAD much in the manner described by Kilpatrick et al. (2001). Nonetheless, it is impossible to interpret the meaning of the unusually complex procedure for the MAD if one is not familiar with the procedure itself, which would seem to explain why those who most successfully interpreted it in this study had also first demonstrated the ability to execute the procedure.

It should be noted that an argument for the importance of procedural knowledge in attaining conceptual knowledge of the MAD does not immediately imply that instruction should consist of practice with procedures before conceptually-oriented problems are posed. The failure of such a teaching pattern to build students' understanding has been well-documented (Mack, 1990; Pesek & Kirshner, 2000; Wearne & Hiebert, 1988). A more plausible implication is that it is worthwhile to help students unpack the unusually complex procedural structure of the MAD while also attending to the conceptual subtleties involved in interpreting it. This sort of careful analysis of the MAD algorithm could provide a strong basis for helping students develop the ability to interpret the conceptual meaning of the measure. If the algorithm is examined in this manner rather than treated merely as a means to the end of computing a "correct" answer, then studying the algorithm may provide a basis for developing conceptual understanding, since the algorithm contains a great deal of information about the meaning of the measure.

Conclusion

The present study begins to fill the void in the literature in the area of teachers' understanding of the MAD. The results reported should help other teacher educators anticipate some of the patterns of thinking they will encounter among the prospective teachers they serve. Additionally, the results may help predict patterns of thinking likely to be present among children, as it has been observed that teachers' knowledge of statistics is often not much different from that of their students (Groth & Bergner, 2006; Jacobbe & Horton, 2010; Leavy, 2006). The similarity between teacher and student knowledge of statistics is likely to be amplified even more in the case of the MAD, as it is simultaneously a new object of study for both teachers and students in many school contexts. As we chart instructional pathways for helping students and teachers attain conceptual understanding of the MAD, the importance of procedural knowledge of the MAD should not be overlooked. Procedural knowledge and conceptual knowledge are deeply intertwined in the case of the MAD, and procedural knowledge may ultimately provide a basis for deep conceptual understanding of the MAD.

References

Aliaga, M., Cobb, G., Cuff, C., Garfield, J., Gould, R., Lock, R., et al (2005). *Guidelines for assessment and instruction in statistics education* (GAISE): *College report.* Retrieved from http://www.amstat.org/education/gaise

Bakker, A. (2004). Reasoning about shape as a pattern in variability. *Statistics Education Research Journal, 3*(2), 64-83. Retrieved from http://www.stat.auckland.ac.nz/~iase/serj/SERJ3%282%29_Bakker.pdf

Baroody, A. J., Feil, Y., & Johnson, A. R. (2007). An alternative reconceptualization of procedural and conceptual knowledge. *Journal for Research in Mathematics Education, 38,* 115-131.

Cobb, G. W., & Moore, D. S. (1997). Mathematics, statistics, and teaching. *American Mathematical Monthly, 104,* 801-823.

Corbin, J., & Strauss, A. (2008). *Basics of qualitative research* (3rd ed.). Thousand Oaks, CA: Sage.

delMas, R., & Liu, Y. (2005). Exploring students' conceptions of the standard deviation. *Statistics Education Research Journal, 4*(1), 55-82.

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.

Garfield, J. B., & Ben-Zvi, D. (2008). *Developing students' statistical reasoning: Connecting research and teaching practice.* New York: Springer.

Groth, R. E. (2012). The role of writing prompts in a statistical knowledge for teaching course. *Mathematics Teacher Educator, 1,*23-40.

Groth, R. E. (2013a). A day in the life of a statistical knowledge for teaching course. *Teaching Statistics, 35*(1), 37-42.

Groth, R. E. (2013b). Characterizing key developmental understandings and pedagogically powerful ideas within a statistical knowledge for teaching framework. *Mathematical Thinking and Learning*

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 (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 1-28). Hillsdale, NJ: Erlbaum.

Hill, H. C., Ball, D. L., & Schilling, S. G. (2008). Unpacking pedagogical content knowledge: Conceptualizing and measuring teachers' topic-specific knowledge of students. *Journal for Research in Mathematics Education, 39,* 372-200.

Jacobbe, T., & Horton, R. M. (2010). Elementary school teachers' com-

prehension of data displays. *Statistics Education Research Journal, 9*(1)*, 27-45. Retrieved from http://www.stat.auckland.ac.nz/~iase/serj/ SERJ9%281&29_Jacobbe_Horton.pdf

Jacobs, J. K., Hiebert, J., Givvin, K., Hollingsworth, H., Garnier, H., & Wearne, D. (2006). Does eighth-grade teaching in the United States align with the NCTM Standards? Results from the TIMSS 1995 and 1999 video studies. *Journal for Research in Mathematics Education, 37,* 5-32.

Kader, G., & Mamer, J. (2008). Statistics in the middle grades: Understanding center and spread. *Mathematics Teaching in the Middle School, 14,* 38-43.

Kilpatrick, J., Swafford, J., & Findell, B. (Eds.). (2001). *Adding it up: Helping children learn mathematics.* Washington, DC: National Academy Press.

Leavy, A. M. (2006). Using data comparison to support a focus on distribution: Examining preservice teachers' understandings of distribution when engaged in statistical inquiry. *Statistics Education Research Journal, 5*(2)*, 89-114. Retrieved from http://www.stat.auckland. ac.nz/~iase/serj/SERJ5(2)_Leavy.pdf

Ma, L. (1999). *Knowing and teaching elementary mathematics.* Mahwah, NJ: Erlbaum.

Mack, N. K. (1990). Learning fractions with understanding: Building on informal knowledge. *Journal for Research in Mathematics Education, 21,* 16-32.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis* (2nd ed.). Thousand Oaks, CA: Sage.

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

Muhr, T., & Friese, S. (2004). *User's Manual for ATLAS.ti 5.0.* Berlin: ATLAS.ti Scientific Software Development GmbH.

National Governors Association for Best Practices & Council of Chief State School Officers. (2010). *Common core state standards for mathematics.* Retrieved from http://www.corestandards.org/

Pesek, D. D., & Kirshner, D. (2000). Interference of instrumental instruction in subsequent relational learning. *Journal for Research in Mathematics Education, 31*, 525-540.

Sánchez, E., da Silva, C., & Coutinho, C. (2011). Teachers' understanding of variation. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics − Challenges for teaching and teacher education* (pp. 211-221). Dordrecht: Springer.

Star, J. R. (2005). Reconceptualizing procedural knowledge. *Journal for Research in Mathematics Education, 36,* 404-411.

Star, J. R. (2007). Foregrounding procedural knowledge. *Journal for Research in Mathematics Education, 38,* 132-135.

Wearne, D., & Hiebert, J. (1988). A cognitive approach to meaningful mathematics instruction: Testing a local theory using decimal numbers. *Journal for Research in Mathematics Education, 19,* 371-384.

Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical inquiry. *International Statistical Review, 67,* 223-265.