Proceedings of the Annual Meeting of the Georgia Association of Mathematics Teacher Educators

Volume 13
Issue 1 *13th Annual Proceedings*

Article 2

2019

Conceptualizing and Interpreting Mean and Median With Future Teachers

Eryn M. Stehr
Georgia Southern University, estehr@georgiasouthern.edu

Ha Nguyen
Georgia Southern University, hnguyen@georgiasouthern.edu

Gregory Chamblee
Georgia Southern University, gchamblee@georgiasouthern.edu

Sharon Taylor
Georgia Southern University, taylors@georgiasouthern.edu

DOI
10.20429/gamte.2019.130102

Conceptualizing and Interpreting Mean and Median

With Future Teachers

Eryn M. Stehr[1], Ha Nguyen[1], Gregory Chamblee[2], and Sharon Taylor[1]
[1] Department of Mathematical Sciences, Georgia Southern University
[2] Department of Middle Grades and Secondary Education, Georgia Southern University

Abstract

Mathematical Education of Teachers II (METII), echoed by the American Statistical Association publication, *Statistical Education of Teachers*, recommended teacher preparation programs support future teachers in developing deep understandings of mean and median, such that middle grades teachers may use them to "summarize, describe, and compare distributions" (Conference Board of Mathematical Sciences, 2012, p. 44; Franklin et al., 2015). *Georgia Standards of Excellence* require statistical reasoning from students beginning as early as 6-7 years old, including interpretation of measures of center and statistical reasoning about best measures of center (Georgia Department of Education, 2015). This level of understanding and interpretation of measures of center, however, has been a persistent struggle for students and their teachers (e.g., Jacobbe & Carvalho, 2011). Jacobbe and Carvalho argued that an over-reliance on computation with little focus on conceptual understanding has created these barriers to statistical reasoning. To impact students' understanding, a starting point is to address teachers' understanding, particularly by supporting conceptual understanding of measures of center in teacher preparation programs (Jacobbe & Carvalho, p. 207). Our research question was: *What conceptual understandings of mean and median do preservice teacher candidates (PSTs) exhibit when presented with a mean and median statistical task?* We present findings from a two-part study, comparing PSTs' responses to a task written to elicit conceptual understandings and statistical reasoning in one semester, with PSTs' responses to a revised task in a second semester, both given at the end of a senior-level Statistics for K-8 Teachers course.

Keywords: statistics, measures of center, mathematics education

Conceptualizing and Interpreting Mean and Median with Future Teachers

Calls for deepening future teachers' understanding of measures of center are supported by recent policy documents. Conference Board of Mathematical Sciences (CBMS, 2012) recommended teacher preparation programs support future teachers in developing deep understandings of mean and median, such that middle grades teachers may use them to "summarize, describe, and compare distributions" (CBMS, 2012, p. 44). This position was echoed by the American Statistical Association publication *Statistical Education of Teachers* (Franklin et al., 2015). *Georgia Standards of Excellence* require statistical reasoning from students as early as first grade, including interpreting measures of center and statistical reasoning about best measures of center (Georgia Department of Education, 2015).

This level of interpreting measures of center, however, has been a persistent struggle for students and their teachers (e.g., Jacobbe & Carvalho, 2011). Jacobbe and Carvalho argued that an over-reliance on computation with little focus on conceptual understanding has created these barriers to statistical reasoning. To impact students' understanding, a starting point is to address teachers' understanding, particularly by supporting conceptual understanding of measures of center in teacher preparation programs (Jacobbe & Carvalho, p. 207).

We asked: *What conceptual understandings of mean and median do preservice teacher candidates (PSTs) exhibit when presented with a mean and median statistical task?*

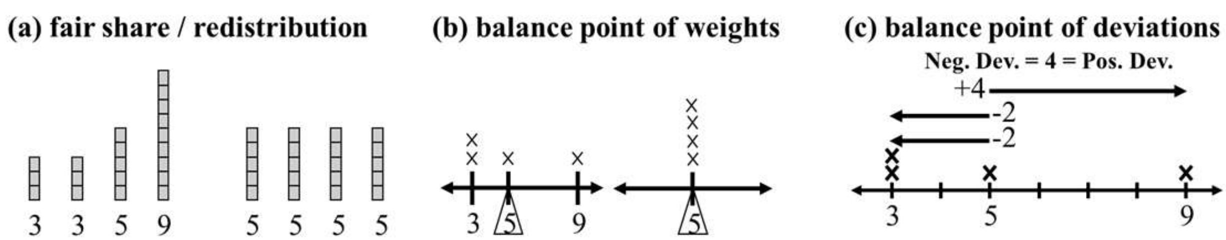
BACKGROUND

The central tendency of a set of data is an attempt to describe collected data by identifying a single value as a reasonable or fair representation of the data set as a whole. That is, a measure of center should summarize data with one number representing all values fairly. To decide which measure accomplishes this fair summary, distribution of data is considered and typically either the arithmetic mean or the median is chosen as the measure of center.

The arithmetic mean is calculated by dividing the sum of all data values by the number of values in the data set. The Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report described levels of conceptual understanding needed to develop robust understanding of the mean: to reach an abstract and robust understanding, students must first understand finding mean as fair sharing, then as a balance of weights, and finally as the point where positive and negative deviations sum to zero (Franklin et al., 2007). To understand mean as *fair sharing* (or a *redistribution* or a *leveling out*) values in a data set, students combine and then redistribute values of all elements, so each is equal (Figure 1a). Preservice teachers (PSTs) experiment by leveling or redistributing physical or visual block towers (Beckmann, 2011).

Figure 1

Arithmetic Mean from Concrete to Abstract

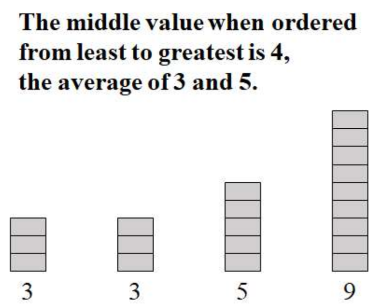


To build on this understanding, focus shifts to *balance point of values* viewed as weights along a rod (Figure 1b). After physically or visually finding the balance point, students begin calculating and summing deviations (distance between data point and mean). When the data point is greater than or less than the mean, the distance is called positive or negative deviation, respectively (Figure 1c). The mean (balance point) is the value where the sums of positive and negative deviations are equal. GAISE argued that supporting students as they move from conceptualizing mean as fair sharing to balancing, and then to the more abstract deviations supports students in later understanding of center and variation (Franklin et al., 2007).

The median of a data set is its middle score (or average of two middle scores) when the data are arranged by magnitude (see Figure 2). Half of the observed values are greater than or equal to the median, and half are less than or equal to.

Figure 2

Median as a Middle Value When Data are Ordered From Least to Greatest



To determine the fairest representation of a data set, students consider the shape of the distribution (i.e., symmetric or asymmetric). Note that the mean of the data in Figure 2 is 5 while the median is 4. If only one value is changed drastically, then the mean would but the median would not. That is, when many data points cluster with only a few very large or very small data points, then the mean will not fairly represent the data but the median will.

METHOD

Participants in the study were 52 senior- and junior-level PSTs enrolled in three sections across two semesters of a *Probability and Statistics for K-8 Teachers* course (Table 1). Measures of center are introduced in a prerequisite course, *Foundations of Data and Geometry*, which emphasizes computing and defining measures of center, rather than applying or interpreting. The *Probability and Statistics for K-8 Teachers* course requires PSTs to apply and interpret these measures across a variety of situations and contexts.

Table 1

Certification Programs of Senior- and Junior-Level Participants

Certification Program	Spring 2018	Spring 2019
Elementary (Early Childhood)	32	1
Middle Grades	3	7
Dual Certification	6	0
Special Education	3	0

During the final session of each *Probability and Statistics for K-8 Teachers* course (taught by the second author), PSTs individually completed a modified version of a 2007 National Assessment of Educational Progress (NAEP) test item given to students 13-14 years old in the United States (National Center for Education Statistics, 2007), using any type of calculator (Figure 3). The original 2007 NAEP item required participants to only answer one question, part (c). Spring 2018 participants completed the original 2007 NAEP item plus two additional items [parts (a) and (b)]. Spring 2019 participants completed a modified version of the Spring 2018 question [italicized].

Figure 3

Study Instrument: Modified NAEP Item

The table below shows the number of customers at Malcolm's Bike Shop for 5 days, as well as the mean (average) and the median number of customers for these 5 days.

Number of Customers at Malcolm's Bike Shop	
Day 1	100
Day 2	87
Day 3	90
Day 4	10
Day 5	91
Mean (average)	75.6
Median	90

a) Interpret the mean. [*in the context of Malcolm's Bike Shop*.]
b) Interpret the median. [*in the context of Malcolm's Bike Shop*.]
c) Which statistic, the mean or the median, best represents the typical number of customers at Malcolm's Bike Shop for these 5 days? Explain your reasoning.

The authors coded items as correct, partially correct, or incorrect for each sub-part separately. Parts (a) and (b) were coded based on each author's own judgment. Part (c) was coded using NAEP item assessment data. The authors met and discussed each sub-part during one meeting. The authors provided their scoring for each item. Scoring discrepancies were discussed as a team. Sub-part answers that did not receive consensus during the first scoring meeting were re-scored individually and a follow-up meeting was held. Consensus scoring was determined at this meeting.

FINDINGS

In Table 2, results for 44 PSTs in Spring 2018 and 8 PSTs with the revised question in Spring 2019 are shown by sub-part. We compared results of the initial and revised tasks, to see if the revisions prompted different responses. Because we added "in the context of Malcolm's Bike Shop" to parts (a) and (b), we made the division between partially correct and incorrect stricter; that is, any responses that did not include context were scored as incorrect. Because of this change, we only report the number of fully correct responses here. We saw changes but cannot definitively conclude if such changes were due to changes in the task or in the PST population.

Table 2

Scoring Results for the Three Sub-Parts of the Task by Semester

Scoring	(a) Mean	(b) Median	(c) Best Statistic
Spring 2018 (n = 44)	11 (25%)	0 (0%)	20 (45.5%)
Spring 2019 (n = 8)	3 (37.5%)	0 (0%)	5 (62.5%)

Task sub-part (a) mean. In scoring PSTs' responses to (a) for correctness, we considered elements of our expected correct response: "*If Malcolm's Bike Shop had the same number of total customers and distributed them evenly among the 5 days, he would have 75.6 customers per day*." We scored PSTs' responses as correct when they included context and when there was some evidence that they understood the *mean* as fair sharing values or as a balance point for weights or deviations. Even in "correct" responses, PSTs often used circular logic, stating *mean* as the *average* rather than supplying conceptual meaning. For example, from Spring 2018, an example of a correct response is: "The average number of customers present on any given day is 75.6 customers." Similarly, a correct response from Spring 2019 is "The mean number of customers at the shop is 75.6 on any given day." As mentioned above, we scored these responses as correct because the PST included context (e.g., customers) along with the phrasing "any given day" which gives a sense of redistributing the customers among the five days. One significant change from Spring 2018 to Spring 2019 centered on re-calculating the mean. In Spring 2018, 12 of the 44 PSTs calculated the mean without connecting to context. For example, "100+ 87+90+10+91=378 [leads to] 378.0/5= 75.6" shows the calculation of the mean even though

75.6 is already included in the statement of the task. In Spring 2019, no PST simply calculated the mean, and all but one of the PSTs connected to the context.

Task sub-part (b) median. Our expected correct response for part (b) was: "Half of the days at Malcolm's Bike Shop had less than (or equal to) 90 customers and half had more than (or equal to) 90 customers." In both semesters, no PST gave a response that we scored as correct because no PST indicated conceptual understanding by indicating customers on half of the days would be above 90 and half below. Similar to findings from part (a), the change from Spring 2018 to Spring 2019 was that over half of the PSTs in 2018 did not connect to context at all, while only one in 2019 did not.

Task sub-part (c) best statistic. Our expected correct response for the last part, identifying the "best statistic" and explaining their reasoning, was: "*The median best represents the typical number of customers of this data because of the outlier of the 10. The number on Day 4 is so drastically different from the other days, it drastically affects the mean.*" An example of a correct PST response is: "Median [is the best statistic] because there is an extreme low value of 10 which pulls the mean down." The wording of this question was not changed from Spring 2018 to Spring 2019, so we compare the scoring of this task between semesters in Table 3.

Table 3

Scoring Results for Sub-Part (c) by Semester

Scoring	Spring 2018 (*n* = 44)	Spring 2019 (*n* = 8)
Correct	20 (45.5%)	5 (62.5%)
Partially Correct	10 (22.7%)	3 (37.5%)
Incorrect	14 (31.8%)	0 (0.0%)
Total:	44 (100%)	8 (100%)

A response we would score as partially correct is: “The median [because] it is a constant sum measure + there are outliers.” Finally, a fully incorrect response might identify the mean as the best statistic, and often included misconceptions. For example: “Mean because it is a bigger range” or “The mean because it most directly reflects the middle of the data.” One change that must be due to the change in PST population was that all students in Spring 2019 gave correct or partially correct responses, while 32% of PSTs in Spring 2018 gave incorrect responses.

CONCLUSION

These findings show students have a basic understanding of the concepts of mean and median, but also demonstrate that PSTs had not yet reached the level of “contextual interpretation” as recommended in the GAISE Report (Franklin et al., 2007). Although 48% of our PSTs (25 of the 52) correctly selected median as the best measure [sub-part (c) question] as compared to 6% of eighth graders in the NAEP results (NCES, 2007), less than half correct finding does not demonstrate mastery of the concepts for most of the PSTs. We also recommend clarifying wording of the question; although we cannot conclude that adding “in the context of…” to each question helped our students, initial results seem promising.

Because research has shown that statistical misconceptions and mistakes appear frequently in classroom teaching (e.g., Groth, 2017), we recommend that measures of center (i.e., mean and median) are major individual concepts to master conceptually, and to use when comparing and contrasting distributions. When adding the complexity of asking PSTs to choose between the two measures with respect to a data set, instructors need to ensure PSTs know how these measures are interrelated and how they differ. Therefore, we recommend additional instructional time be spent with PSTs applying mean and median to varying levels of complex contextual tasks that require PSTs to demonstrate conceptual understanding of these constructs, and to notice similarities and differences.

References

Estrada, A., Batanero, C., Fortuny, J. M., & Díaz, C. (2006). A structural study of future teachers' attitudes towards statistics. In Bosch, M. (Ed.). *Proceedings of the Fourth Congress of the European Society for Research in Mathematics Education* (pp. 508-517). Sant Feliu de Guíxols, Spain: FUNDEMI IQS – Universitat Ramon Llull and ERME.

Beckmann, S. (2011). *Mathematics for elementary school teachers (3rd ed)*. Boston, MA: Pearson/Addison-Wesley.

Conference Board of the Mathematical Sciences. (2012). *The mathematical education of teachers II.* Providence, RI, and Washington, DC: American Mathematical Society and Mathematical Association of America.

Franklin, C., Bargagliotti, A., Case, C., Kader, G., Scheaffer, R., & Spangler, D. (2015). *Statistical education of teachers*. Alexandria, VA: ASA.

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.

Georgia Department of Education. (2015). *Mathematics Georgia standards of excellence*. Atlanta, GA: Author. Retrieved from https://www.georgiastandards.org/Georgia-Standards/Pages/Math.aspx

Groth, R. E. (2017). Developing statistical knowledge for teaching during design-based research. *Statistics Education Research Journal, 16(*2), 376-396.

Jacobbe, T. (2012). Elementary school teachers' understanding of the mean and median. *International Journal of Science and Mathematics Education, 10(*5), 1143-1161.

Jacobbe, T., & Carvalho, C. (2011). Teachers' understanding of averages. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics: Challenges for teaching and teacher education* (pp. 199-209). New York: Springer.

National Center for Education Statistics. (2007). *The nation's report card: Mathematics 2007.* Washington, DC: Institute of Education Sciences, US Department of Education. Retrieved from http://nces.edu.gov/nationsreportcard/itmrlsx/