Developing Proficiency in Grade 6 Common Core Statistics

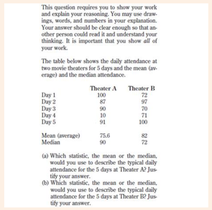
Kristen Kent & Ebony Hitch Mentor: Dr. Randall Groth Salisbury University

Introduction

National Assessment of Education Progress (NAEP) results indicate that middle school students have some difficulty finding the mean and median. Students at all grade levels have considerable difficulty knowing how to use statistical measures of center appropriately (Zawojewski & Shaughnessy, 2000).

The purpose of the study was to explore and develop students' thinking about graphical representations of data and finding appropriate measures of center. For example, in the problem below

Research question: How can students' proficiency in regard to Grade 6 Common Core Mathematics Standards about statistical measures of center be developed?.



This question requires you to show your work and explain your reasoning. You may use drawings, words, and numbers in your explanation. Your answer should be clear enough so that another person could read it and understand your thinking. It is important that you show all of your work.

The table below shows the daily attendance at two movie theaters for 5 days and the mean (average) and the median attendance.

	Theater A	Theater B
Day 1	100	72
Day 2	67	97
Day 3	90	70
Day 4	10	71
Day 5	91	100
Mean (average)	75.6	82
Median	90	72

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

Reference: Zawojewski, J.S., & Shaugnessy, J.M. (2000). Mean and Median: Are they really so easy? *Mathematics Teaching in the Middle School*, 5(7), 436-440

Theoretical Framework

We used the *Adding it Up* framework (National Research Council, 2001, p. 116) to conceptualize mathematical proficiency. It includes the following five strands:
Conceptual Understanding – comprehension of mathematical concepts, operations, and relations
Procedural fluency – skill in carrying out procedures flexibly, accurately, efficiently, and appropriately
Strategic Competence – ability to formulate, represent, and solve mathematical problems
Adaptive reasoning – capacity for logical thought, reflection, and explanation, and justification
Productive disposition – habitual inclination to see mathematics as sensible, useful, and worthwhile, coupled with a belief in diligence and one's own efficiency.

Additional Guiding Concepts from the Literature

Statistical investigation cycle: Groth & Bargagliotti (2012) explained how to engage all students in four processes of statistical investigation using the Common Core by formulating questions, collecting data, analyzing data and interpreting results. To the extent possible, students should experience all phases of statistical investigation to understand the nature of statistics in practice.

Conceptual understanding of measures of center: It is important for students to understand the mean and median as concepts and not just as computations (Bremigan, 2003). In addition to finding, using, and interpreting measures of center, we focused on helping students understand the mean's relationship to other measures of center, such as median and mode.

References
Bremigan, E. G. (2003). Developing a Meaningful Understanding of the Mean. *Mathematics Teaching in the Middle School*, 9(1), 22-26.
Common Core Standards Writing Team. (2011). *Progression for the Common Core State Standards for Mathematics (draft),* 6-8, Statistics and Probability. Retrieved from http://commoncoretools.files.wordpress.com/2011/12/ccss_progression_sp_68_2011_1226_bis.pdf.
Groth, R. E., & Bargagliotti, A. E. (2012). GAISEing into the Common Core of Statistics. *Mathematics Teaching in the Middle School*, 18(1), 38-45.
National Research Council. (2001) *Adding it Up: Helping Children Learn Mathematics*. J. Kilpatrick, J. Swafford, and B. Findell (Eds.). Mathematics Learning Study Committee, Center for Education, Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

CCSSM Learning Progressions for Statistics

The Common Core State Standards Writing team (2011) described key transitions and competencies in learning statistics in accordance with the Common Core State Standards in a learning progressions document. Key ideas from the progressions relevant to the instructional sessions for students were:

Statistical investigations begin with a question, and students now see that answers to such questions always involve variability in the data collected to answer them (p. 4).

Working with counts or measurements, students display data with the dot plots (sometimes called line plots) that they used in earlier grades (p. 4)

A major focus of Grade 6 is characterization of data distributions by measures of center and spread. To be useful, center and spread must have well-defined numerical descriptions that are commonly understood by those using the results of a statistical investigation. The simpler ones to calculate and interpret are those based on counting. In that spirit, center is measured by the median (p. 4).

Students use their knowledge of division, fractions, and decimals in computing a new measure of center—the arithmetic mean, often simply called the mean (p. 5).

Students also learn some of the subtleties of working with the mean, such as its sensitivity to changes in data values and its tendency to be pulled toward an extreme value, much more so than the median. Students gain experience in deciding whether the mean or the median is the better measure of center in the context of the question posed (p. 5)..

Reference:
National Research Council. (2001) *Adding it Up: Helping Children Learn Mathematics.* J. Kilpatrick, J. Swafford, and B. Findell (Eds.). Mathematics Learning Study Committee, Center for Education, Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

Methodology: Participants & Procedures

Over the course of ten weeks, we worked with four students completing fifth grade. We had a 100% participation rate with seven weekly one-hour sessions in addition to our pre and post assessment interviews. For the privacy of the students, the following pseudonyms will be used: Cody, Flynn, Millie, Giselle

Common Core State Standards (CCSS) Instructional Goals (National Governor's Association for Best Practices & Council of Chief State School Officers, 2010, p. 45):

Understand that a set of data collected to answer a statistical question has a distribution which can be described by its center, spread, and overall shape. (CCSS.MATH.CONTENT.6.SP.A.2)

Recognize that a measure of center for a numerical data set summarizes all of its values with a single number, while a measure of variation describes how its values vary with a single number. (CCSS.MATH.CONTENT.6.SP.A.3).

Display numerical data in plots on a number line, including dot plots, histograms, and box plots. (CCSS.MATH.CONTENT.6.SP.B.4).

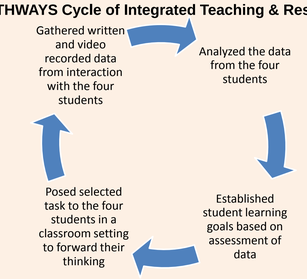
Giving quantitative measures of center (median and/or mean) and variability (interquartile range and/or mean absolute deviation), as well as describing any overall pattern and any striking deviations from the overall pattern with reference to the context in which the data were gathered. (CCSS.MATH.CONTENT.6.SP.B.5.C)

Relating the choice of measures of center and variability to the shape of the data distribution and the context in which the data were gathered. (CCSS.MATH.CONTENT.6.SP.B.5.D)

Reference:
National Governor's Association for Best Practices & Council of Chief State School Officers. (2010). Common Core State Standards for Mathematics. Retrieved from http://www.corestandards.org/

Methodology: Data Gathering & Analysis

PATHWAYS Cycle of Integrated Teaching & Research



During each lesson, two cameras were set up to record the entire hour session. Following the lesson, we would playback the video and transcribe each word spoken, as well as any emotions and movements throughout the lesson. We would then use these transcriptions to find strengths and weaknesses in students' learning in terms of the 5 Strands of Mathematical Proficiency, and then make data-based conjectures about how to foster students' learning along each strand. These conjectures became the basis for developing the following week's lesson.

Empirical Teaching and Learning Trajectory:

Initial Assessment Results (Week 1)

Overall most of the students lacked conceptual understanding when it came to finding typical values.

Millie lacked *conceptual understanding* when representing data and trying to interpret data from dot plots. She had strong *strategic competence* when it comes to comparing statistical measures. In the excerpt below you will see that Millie was able to choose the median as the best representation because most of the numbers were closer to it.

EH: Which statistic, the mean or the median would you use to describe the typical daily attendance for the 5 day at Theater A?
Millie: Median
EH: Why did you choose the median?
Millie: Because Theater A has the most higher numbers and the mean has a low number and the median has the same kind of range as the numbers in Theater B.

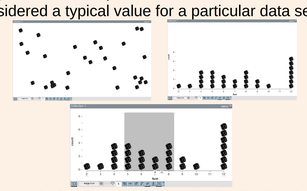
Cody lacks *conceptual understanding* when he is not able to understand which statistic represents the typical daily attendance for the movie theaters.

Giselle had strong *procedural fluency* in regard to making accurate dot plots and finding statistical measures. She lacked some *conceptual understanding* in regard to comparing and finding distributions between graphs.

Flynn lacked some *adaptive reasoning* throughout most of the interview since he would often use knowledge of the context of a problem alone without referring to the data given (e.g., for one task, he answered a question about typical puppy weight by referring to his past experiences but not the data). His lack of *procedural fluency* with basic math skills appeared to have contributed to his difficulties.

Displaying Data (Week 2, 3, 4)

These lessons focused on having the students generate data from rolling two dice in pairs and represent data using dot plots. After becoming familiar with displaying data graphically, they were able to organize and compare multiple data sets within a single graph using TinkerPlots (Konold 2005). This encouraged them to identify a middle clump of the data and think about what is considered a typical value for a particular data set.



Students quickly discovered a method to finding the middle data value (median) by crossing off values from each side of the graph, which contributed to their *procedural fluency*. They were also able to use descriptive words to define the graph's behavior, such as hole, gap, hill, and cluster. Although most of the students were able to find the middle clump accurately, some students still described the middle using the values on the number line, not the actual data values. Overall, students developed both conceptual understanding and *strategic competence* as they were able to represent data graphically using a dot plot and identify where most of the data values congregated on the graph.

Reference: Konold, C., & Miller, C. D. (2005). TinkerPlots: Dynamic data exploration. [Computer software] Emeryville, CA: Key Curriculum Press.

Understanding Mean (Week 5, 6)

As students gained experience with dot plots and comparing data, we started discussing differences in the shapes of graphs and how each of the statistical measures (mean, median, and mode) are affected when the data shapes the graph differently. This generated discussion on what the mean is and how it is affected with different data sets. In particular, students began understanding mean as a number that "evens out" or "balances" a distribution as they began using snap cubes as data values for number of family members in a particular household. By redistributing the snap cubes (or family members) they could easily see how the mean represented a "fair share" for the data set.

Students worked on contrasting the snap cube representation above with the dot plot representation they had learned earlier. Giselle was also able to identify the mode as the highest stack of data values in a dot plot. This generated class discussion about what is most typical and if the mode always represented what was most typical. In particular, Cody demonstrated multiple gains in *conceptual understanding* as he accurately identified the middle clump and recognized that the mode is not always the best representation for the graph. Students gained *strategic competence* as they were able to explain their reasoning through graphs and showing others their methods that supported their reasoning. Flynn also used the snap cubes to describe the behavior of the mean, but still had some confusion with the middle values being middle data values as opposed to the middle of the number line.

Measures of Center (Week 7, 8)

As students became more familiar with what the mean represents, we presented them with skewed data sets and asked them to decide whether or not the mean was a good representation of the overall data set. This initiated discussion on whether the mean, median, or mode did a better job at representing the data. Students quickly discovered various reasons for choosing certain measures of center to represent different data sets. By the last week, we were able to review majority of the topics covered and make inferences about how and why it is important to use different measures of center to describe data.

For example, when 24 Starburst candy were distributed unevenly amongst the four students and two teachers, students were asked to find the average or typical number of candies that each person received. The initial candy distribution was the following: **1, 1, 1, 1, 1, and 19**. Two students described the mode as the average; Millie tried to find the average by finding the middle between the lowest and highest data values while Cody explained that he thought the average was the mean and redistributed the candy evenly to show a "fair share". Students quickly realized that the **mean is not always a good representation of the typical value**. Overall, students were accurately able to identify the mean, median, and mode and how each value described the data.

By the last week, we were able to review majority of the topics covered and make inferences about how and why it is important to use different measures of center to describe data. During the final lesson, students analyzed a data set showing the salaries of individuals in a small town: **$0, $0, $0, $0, $0, $0, $0, $0, $200, $200, $200, $200, $200, $200, $30,600**

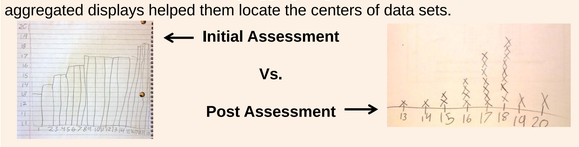
Reference:
Lappan, G., Fey, J.T., Fitzgerald, W.M., Friel, S.N., & Philips, E.D. (2004). *Data About Us*. New York: Pearson.

Post-Assessment Results (Week 9)

Overall students gained *conceptual understanding* when it came to finding the best statistical measure to represent a typical value. In the initial interviews, most students were not able to provide a rationale for selecting the mean or the median to represent the typical value for a data set. In the final interviews, most students choose between the mean and median by looking for the measure around which the data clustered.

Initial Assessment	Post Assessment
KK: So do you think the mean is a better representation of the numbers or do you think the median is a better representation of the numbers? Flynn: Median KK: Why do you say median? **Flynn: Because you got to put them least to greater.**	EH: Which statistic, the mean or the median would you use to describe the typical daily attendance for the 5 day at Theater A? Flynn: Median EH: Why did you choose the median? **Flynn: Because most of the numbers are close to 90.**

Students gained *procedural fluency* and *strategic competence* in selecting and constructing data displays. In the initial interviews, students chose to represent data by showing each data value with a separate bar. In the final interviews, students represented the same data sets by aggregating the data in dot plots. These aggregated displays helped them locate the centers of data sets.



Some students gained *productive disposition*. Giselle, for example, did not volunteer her thinking during lessons at the outset of the summer. During the last two instructional sessions, she become a consistent contributor to class discussion. In one session, she explained the algorithm for the mean in terms of the snap cube model used to build the idea of mean as a fair share.

Reflection

After researching student thinking and understanding over a ten-week period, we have learned that achieving every CCSSM Standard for Grade 6 Statistics is very challenging. During this study, we were able to help students begin to reason conceptually about measures of center, but did not have time to delve into formal measures of variability, which are also prescribed in the sixth-grade Common Core. There is also additional work students will need to do as they further develop understanding of center required in the Common Core. In particular, it was difficult teaching the concept of connecting how changing some of the data could affect the measures of center. Another challenge included some students switching back and forth between dot plots and case value bars, especially seen in our lessons on Understanding Mean. In order to lessen the confusion between the two, we would ask probing questions such as "What does each dot represent?" and "What are you trying to graph?" in order to probe students to think about what is actually happening with the graph or data. Because of the conceptual complexity of learning measures of center as required in the Common Core, it may be beneficial for schools to begin to develop these ideas before sixth grade, so that students also have adequate time to engage meaningfully in the study of formal measures of variability. The recommendation to begin the study of these concepts before sixth grade resonates with recommendations found in documents such as Guidelines for Assessment and Instruction in Statistics Education (Franklin et. al., 2007) and Principles and Standards for School Mathematics (National Council of Teachers of Mathematics, 2000).

References:
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.
National Council of Teachers of Mathematics. (2000). *Principles and Standards for School Mathematics*. Reston, VA: Author.