Chapter 9: Learning to Reason about Center

A statistician sees group features such as the mean and median as indicators of stable properties of a variable system—properties that become evident only in the aggregate. (Konold & Pollatsek, 2002, p. 262)

Snapshot of a Research-based Activity on Center

Pairs of Students are given a set of 10 small Post-It® notes and a number line that goes from 17 to 25. They are asked to think about a group of college students and their ages. They are told that the mean age is 21, and are asked to construct a dot plot, using the Post-It® notes, of a distribution with a mean of 21. Most of them quickly figure out that they can stack all 10 Post-It® notes at the 21 point. But then they are told that one of the values is 24, so they have to figure out where the other values are on the number line. Next they are told to move one Post-It® note to 17, and arrange the rest of the Post-It® notes so the mean is still 21. Finally, they are asked to move all of the Post-It® notes so that none are at 21 but the mean age is 21.

Students then move to making dot plots of 10 points that have the same mean of 21, and then are asked to draw deviations from the mean for their graphs. They consider how the deviations balance each other out, so if that if one value is moved so that it has a deviation of -3, another value must be moved to have a deviation of +3. This leads to a discussion of what a mean “means” in terms of these deviations all cancelling each other out to be zero.

Rationale for this Activity

Although most student have learned how to calculate means before entering their statistics course, few understand what a mean really is or what it tells us about data. Students often have difficulty understanding how the mean and median differ and why they behave differently (e.g., in the presence of outliers). This activity helps students to
build a conceptual understanding of what the mean and median actually mean and how they are affected by the different values in a data set. This lesson also introduces the idea of deviation early in the course, as a way to understand the idea of a mean and what it represents. This deviation idea is revisited when learning about variability and the standard deviation, and again when considering residuals in a regression analysis. This activity helps students develop a conceptual understanding of the mean rather than a procedural one. By having the students physically manipulate data values on a number line they are better able to see and reason about the idea of deviation from the mean and the balancing of these deviations.

**The Importance of Understanding Center**

*The idea of data as a mixture of signal and noise is perhaps the most fundamental concept in statistics.* (Konold & Pollatsek, 2002, p. 259)

Understanding the idea of center of a distribution of data as a signal amidst noise (variation) is a key component in understanding the concept of distribution, and is essential for interpreting data graphs and analyses. While students develop informal ideas of center in the earlier units as they graph and describe distributions of data, they later encounter the idea of center more formally as they learn about different measure of center, how to compute them, what information they provide, and how we use them. However, it is impossible to consider center without also considering spread, as both ideas are needed to find meaning in analyzing data.

**The Place of Center in the Curriculum**

Traditional textbook first introduce center, then introduce spread, and then move on to the next topic. However it may be more helpful to study these topics together because they are so interrelated (Konold & Pollatsek, 2002; Shaughnessy, 1997).

It is hard to imagine a situation where one would summarize a data set using only a measure of center without talking about the spread of the data as well, or how much variability there is around that measure of center. When comparing groups or making
Inferences we need to examine center and spread together: the signal, and the noise around the signal. While these ideas are introduced in early units on exploring data, these concepts reappear when looking at theoretical models such as the normal distribution and sampling distributions. Later on, the ideas of center (and spread) are revisited when making statistical inferences about samples of data.

**Review of the Literature Related to Reasoning about Center**

*Computational rules not only does not imply any real understanding of the basic underlying concept, but may actually inhibit the acquisition of more adequate (relational) understanding.* (Pollatsek, Lima, & Well, 1981, p. 202)

**Understanding Means**

How students understand ideas of center has been of central interest in the research literature. Research on the concept of average or mean was at first the most common topic studied on learning statistics the school level (see Konold & Higgins, 2003; Shaughnessy, 1992, 2003). The studies suggested that the concept of the average is quiet difficult to understand by children, college students and even elementary school preservice and in-service teachers (Russell, 1990; Groth & Bergner, 2006).

Early research typically focused on the single idea of center rather than looking at the interrelated concepts of center and spread, and on procedural understanding. These studies focused primarily on the mean, either as a simple average of a single data set or a weighted mean. An early interest was on students understanding the mean as a balance model (e.g., Hardiman, Well, & Pollatsek, 1984; Strauss, 1987), which is a common method taught in a statistics course. A balance model illustrates how values are placed on a balance beam at distances from the mean so that the deviations from the mean are equal. Hardiman et al. (1984) tested whether improving students’ knowledge of such balance rules through experience with a balance beam promoted understanding of the mean. Forty-eight college students enrolled in psychology classes participated in the study which involved pretest, training, and posttest of paper and pencil items. Students
who were given the balance training performed significantly better on the posttest problems.

Other studies identified several characteristics of the mean and then examined students’ understanding of these characteristics (Goodchild, 1988; Mevarech, 1983; Strauss & Bichler, 1988; Leon & Zawojewski, 1993). Mevarech (1983), for example, found that high school students made mistakes in solving problems about means because they believed that means have the same properties as simple numbers, and that it is helpful to provide students corrective-feedback instruction as they solve problems involving reasoning about the mean. Strauss and Bichler (1988) found that fourth- through eighth-grade students had a difficult time understanding seven properties of the mean. Leon and Zawojewski (1993) looked at school and college students’ understanding of four components of the mean. Using different testing formats, they found that some properties of the mean are better understood than others. The two properties, “the mean is a data point located between the extreme values of a distribution”, and “the sum of the deviations about the mean equals zero” were better understood by students in this study than the two properties, “when the mean is calculated, any value of zero must be taken into account”, and “the average value is representative of the values that were averaged.”

Gal, Rothschild and Wagner (1989, 1990) found that sixth-grade students are generally unable to use the mean to compare two different-sized sets of data. Later work showed that students’ had difficulty working backward from a mean to a data set that could produce such a mean (Cai, 1998). Study by Mokros and Russell (1995) expanded on this task by having students manipulate data values to produce a given mean and studying how students reasoned during this process.

Earlier research has also concentrated on understanding weighted means. Pollatsek et al. (1981) reported data from interviews of college students indicating difficulties they had in understanding the need to weight data in computing a mean. While mathematically sophisticated college students can easily compute the mean of a group of numbers, this study indicated that a surprisingly large proportion of these students do not understand the concept of the weighted mean which is a concept that they often encounter (e.g.,
grade point averages). When asked to calculate a mean in a context that required a weighted mean, most subjects answered with the simple, or unweighted mean of the two means given in the problem, even though these two means were from different-sized groups of scores. Callingham (1997) found that the same problem in a study of pre-service and in-service teachers. As a result of their study, Pollatsek et al. (1981) wrote that “for many students dealing with the mean is a computational rather than a conceptual act” (p. 191). They concluded that knowledge of “computational rules not only does not imply any real understanding of the basic underlying concept, but may actually inhibit the acquisition of more adequate (relational) understanding” (p. 202).

What students remember about the mean? In general, it appears that many students who complete college statistics classes are unable to understand the idea of the mean. Mathews and Clark (2003) analyzed audio-taped clinical interviews with eight college freshmen immediately after they completed an elementary statistics course with a grade of “A”. The point of these interviews was not to see how quickly isolated facts could be recalled, nor was the point to see how little students remember. Rather, the goal was to determine as precisely as possible the conceptions of mean, standard deviation and the Central Limit Theorem which the most successful students had shortly after having completed a statistics course. The results are alarming since these top students demonstrated a lack of understanding of the mean, and could only state how to find it, arithmetically. Interviewing along the same lines a larger (n=17) and more diverse sample of college students from four distinct campuses, Clark, Karuat, Mathews, and Wimbish (2003) found overall the same disappointing results. These researchers call therefore for pedagogical reform that will dis-equilibrate the process image of statistical concepts that students bring with them to college in order to enable them to encapsulate the process of statistical concepts into objects that are workable entities (Sfard, 1991). For example, they recommend creating situations in which students have to determine and reflect which measure of center is more appropriate.

**Understanding Medians**
Difficulties in determining the medians of data sets have also been documented by research. Elementary school teachers have difficulty determining the medians of data sets presented graphically (Bright & Friel, 1998). Only about one-third of twelfth grade students in the United States who took the NAEP test were able to determine the median when given a set of unordered data (Zawojewski & Shaughnessy, 2000).

Selecting an Appropriate Measure of Center

Another focus of research has been on the challenge of choosing an appropriate measure of center to represent a data set. The National Assessment of Education Progress (NAEP) data confirm that school students frequently make poor choices in selecting measures of center to describe data sets (Zawojewski & Shaughnessy, 2000). Choosing an appropriate measure of center was also a challenge for students enrolled in an Advanced Placement high school statistics course (Groth, 2002). Similar results were found by Callingham (1997) who administered an item containing a data set structured so that the median would be a better indicator of center than the mean, to a group of pre-service and in-service teachers. Callingham report that most of them calculated the mean instead of the more appropriate median.

In a study on statistical reasoning about comparing and contrasting mean, median, and mode of preservice elementary school teachers, Groth and Bergner (2006) described four levels. Their study illustrated that attaining a deep understanding of these seemingly easy statistical concepts is a nontrivial matter and that there are complex conceptual and procedural ideas that need to be carefully developed.

Measures of Center as Typical Values

The typical value interpretation of the arithmetic mean has received a great deal of attention in curriculum materials and in research literature (Konold & Pollatsek, 2002). The following is an example of a problem set in a typical value context:

The numbers of comments made by 8 students during a class period were 0, 5, 2, 22, 3, 2, 1, and 2. What was the typical number of comments made that day? (Konold & Pollatsek, 2002, p. 268).
Several studies have provided insights about students’ thinking in regard to typical value problems. Mokros and Russell (1995) studied the characteristics of fourth through eighth graders’ constructions of “average” as a representative number summarizing a data set. Twenty one students were interviewed, using series of open-ended problems that called on children to construct their own notion of mathematical representativeness. They reported that students may respond to typical value problems by: (i) locating the most frequently occurring value; (ii) executing an algorithm; (iii) examining the data and giving a reasonable estimate; (iv) locating the midpoint of the data; or (v) looking for a point of balance within the data set. These approaches illustrate the ways in which school students are (or are not) developing useful, general definitions for the statistical concept of average, even after they have mastered the algorithm for the mean.

In an investigation of the development of school students’ thinking in regard to typical value problems Watson and Moritz (1999) placed a developmental structure on the categories of thinking documented by Mokros and Russell (1995). They found, “A primitive median concept as ‘middle’ or mode concept as ‘same as others’ is usually acquired well before students are introduced to the arithmetic mean” (p. 35). Jones, Thornton, Langrall, Mooney, Perry and Putt (2000) and Mooney (2002) found that the ability to be thoughtful and critical about applying formal measures to typical value problems marks a relatively sophisticated level of statistical reasoning.

*Measures of center as “signals in noisy processes”*

Students should be given more opportunities to work with statistical problems set in contexts that involve searching for “signals in noisy processes.” The following item is an example of a data analysis problem that involves detecting a signal in a noisy process:

A small object was weighed on the same scale separately by nine students in a science class. The weights (in grams) recorded by each student were 6.2, 6.0, 6.0, 15.3, 6.1, 6.3, 6.2, 6.15, 6.2. What would you give as the best estimate of the actual weight of this object? (Konold & Pollatsek, 2002, p. 268).
In the case of the repeated measures problem above, the arithmetic mean of the weights that are bunched closely together could be viewed as a signal that estimates the true weight of the object. The measurement of the object can be viewed as a noisy process that contains variation stemming from various possible sources. Konold and Pollatsek (2002) acknowledge the possible cognitive complexity in using repeated measurement problems with students, pointing out that the mean as a reliable indicator of signal was not universally accepted by scientists during the early development of the discipline of statistics (Stigler, 1986). Hence, they call for more research on students’ thinking in such contexts in order to help advise instruction.

Patterns of thinking about average in different contexts were investigated by Groth (2005) who studied fifteen high school students. He used problems set in two different contexts: determining the typical value within a set of incomes and determining an average set in a signal-versus-noise context. Analysis of the problem-solving clinical interview sessions showed that some students attempted to incorporate formal measures, while others used informal estimating strategies. Students displayed varying amounts of attention to both data and context in formulating responses to both problems. Groth pointed out the need for teachers to be conscious of building students’ statistical intuitions about data and context and informal estimates of center and connecting them to formal measures without implying that the formal measures should replace intuition.

Using the History of Measures of Center to Suggest Instruction

The history of statistics can be a source of inspiration for instructional design. Based on systematically selected historical examples, Bakker and Gravemeijer (2006) formulate hypotheses about how students in grades 7 and 8 (12–14-years old) could be supported in learning to reason with mean and median. The following ideas (hypotheses) stemming from the historical phenomenology were found to be most fruitful for helping young children understand center.

H1. Estimation of large numbers can challenge students to develop and use intuitive notions of mean.
H4. Students may use the midrange as a precursor to more advanced notions of average.

H5. Repeated measurement may be a useful instructional activity for developing understanding of the mean (cf., Petrosino, Lehrer, & Schauble, 2003).

H8. To support students’ understanding of the median it is helpful to let them visually estimate the median in a dot plot and look for a value for which the areas on the left and right are the same.

H11. Skewed distributions can be used to make the usefulness of the median a topic of discussion.

Such a historical study can help to “unpack” and distinguish different aspects and levels of understanding of statistical concepts and help instructional designers to look through the eyes of students. Note that some of these hypotheses are in accordance with the results of the research studies described above.

**Implications of the Research: Teaching Students to Reason about Center**

What has been striking over 25 years of research is the difficulty encountered by students of all ages and teachers in understanding concepts related to center. Although students may be able to compute simple arithmetic means, they need help in understanding what means actually mean. Activities can help students develop meaningful models such as balancing of data values by manipulating deviations from the mean to sum to zero.

The research suggests that careful attention be paid to developing the concepts of measures of center, focusing on mean and median rather than mode. These ideas should be first introduced informally as students are asked to estimate and reason about typical value for data sets, both large and small, prior to formally studying these topics in a unit on measure of center. Students may be asked to make and test conjectures about typical values using real data sets. The research also suggests that students have opportunities to explore the characteristics of the mean and median and how they are affected by different types of data sets and distributions. Developing an understanding of deviation may be an
important part of understanding the mean as a balance point, so activities helping students to see and reason about deviations may help them better understand the mean. The literature suggests both visual, interactive activates as well as explorations with real data utilizing technology to produce measures of center, especially for data sets where values are changed (e.g., outliers are added or removed). Finally, the idea of the center as a signal in a noisy process should be developed, examining trends in repeated measurements. This also suggests that ideas of center be introduced along with ideas of spread or variability, and that these ideas are repeatedly connected as students explore and interpret data.

Progression of Ideas: Connecting Research to Teaching

Introduction to the Sequence of Activities to Develop Reasoning about Center

The idea of typical value as a summary measure of data set, shown graphically, is first introduced in early units whenever students make or examine graphs of numerical data. While students may intuitively look at the mode or peak of a graph or look at the middle value on the horizontal axis, they can be guided to think about the mean and median as typical values by looking at different graphs where mode or middle scale value do not seem to represent good ‘typical’ values. This will help motivate the need for examining different measures of center and when to use them. These informal examinations and estimates should include estimates of spread of the data as well, as students are asked to respond to questions such “what is a typical value for these data” and “how spread out are the data?”

When formal measures of center are introduced, students are guided to explore their properties using physical and then computer manipulations of data. Properties of the mean and median can be explored and examined in this way. It is helpful for students to actually work backwards, starting with a given value of mean or median to reason about how different data sets may be constructed and altered to produce those given values. This can be done first for mean and then for median. Students are then asked to make conjectures about what typical values they might find for different types of variable, taking into account the shapes and characteristics of graphs of these variables. These
conjectures can then be tested using real data and technology, and discussions can examine which measures are more useful summaries for each variable and why.

When students begin to study formal measures of variability, they see the relationship between mean and standard deviation, and between medial and Interquartile Range, and how it makes sense to use these pairings when summarizing different types of distributions (e.g., means and standard deviations for symmetric distributions, medians and IQR for very skewed distributions and distributions with outliers). The idea of examining center at the same time as variability as a way to compare groups is then encountered as students learn about and compare boxplots. When the Normal distribution is introduced students will see that the mean has special properties and use in relation with standard deviations and z-scores.

The mean is again examined when learning about samples and how the mean stabilizes as sample size gets very large, and the role of the mean in the Central Limit Theorem. As students move from sampling to statistical inference, they again encounter the mean, distinguishing between using the mean in an inference based on a large sample from using the mean as a summary measure of a single data set (when a median might be a better typical value given the shape of the distribution). The measures of center are also encountered in the unit on covariation when students look at trends by examining medians of sequential boxplots, and later as centers of distributions of the two variables.

Table 1: Sequence of activities to develop reasoning about center.

<table>
<thead>
<tr>
<th>Milestones: Ideas and Concepts</th>
<th>Suggested Activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>INFORMAL IDEAS PRIOR TO FORMAL STUDY OF CENTER</td>
<td></td>
</tr>
<tr>
<td>• Idea of center as a typical or representative value for a graph of a variable (e.g., dot).</td>
<td>• Distinguishing Distributions (Lesson 1, Distribution Unit, Chapter 8)</td>
</tr>
<tr>
<td>• The mean as somewhere in between</td>
<td>• What does the Mean Mean Activity</td>
</tr>
</tbody>
</table>
the highest and lowest value but not necessarily the middle value or the midpoint of the horizontal scale.  
(Lesson 1: “Reasoning about Measures of Center”)

### FORMAL IDEAS OF CENTER

- **Properties of the mean as a balance point and the value for which all deviations from that value sum to zero.**  
  - What does the Mean Mean Activity (Lesson 1)

- **How the mean is affected by extreme values.**  
  - What does the Mean Mean Activity (Lesson 1)

- **The median as the middle value in a data set.**  
  - What does the Median Mean Activity (Lesson 1)

- **Properties of the median: under what conditions it changes or stays the same.**  
  - What does the Median Mean Activity (Lesson 1)

- **Comparing properties of the mean and median.**  
  - Means and Medians Activity (Lesson 1)

- **The idea of a typical value.**  
  - What is Typical Activity (Lesson 2: “Choosing Appropriate Measures”)

- **Understanding why and how to use appropriate measures of center for a sample of data for a particular variable.**  
  - Choosing an Appropriate Measure of Center Activity (Lesson 2)

### BUILDING ON FORMAL IDEAS OF CENTER IN SUBSEQUENT TOPICS
• How center and spread are used together to compare groups.

• Activities in Lessons 1, 2, 3 and 4, Comparing Groups Unit (Chapter 11)

• The idea and role of mean in normal distribution.

• What is Normal Activity (Lesson 3, Statistical Models Unit, Chapter 7)

• Recognize stability of measures of center as sample size increases. When sample grows, see how measure of center predict center of larger population, and how it stabilizes (varies less).

• Sampling activities in Lessons 1, 2, and 3, Samples and Sampling Unit (Chapter 12)

• Role of mean in making inferences.

• Activities in Lessons 1, 2, 3, 4 and 5 (Statistical Inference Unit, Chapter 13)

• Role of mean in bivariate distribution.

❖ An activity involving fitting and interpreting the regression line to bivariate data. (The symbol ❖ indicates that this activity is not included in these lessons.)

Introduction to the Lessons

There are two lessons on reasoning about measures of center. They begin with the physical activity described earlier where students manipulate Post-It® notes on a number line to develop an understanding of mean, and then median. Students use a Fathom demo to contrast how the mean and median behave for different types of data sets. Students make and test conjectures about typical values, testing these using software to generate graphs and statistics. The last activity has students compare features and uses of different measures of center when summarizing sample data.
Lesson 1: Reasoning about Measures of Center

While students have heard of means and medians before they enter an introductory high school or college statistics course, this lesson helps them develop a conceptual understanding of the mean and median. There are three parts to the lesson: an activity where students move dots on a plot to explore properties of the mean, a similar activity with the median, and then Fathom demos to further illustrate the properties of these concepts. Student learning goals for this lesson include:

1. Develop a conceptual understanding of the mean.
2. Understand the idea of deviations (differences from the mean) and how they balance out to zero.
3. Understand how these deviations cause the mean to be influenced by extreme values.
4. Develop an understanding of the median as a middle value that is resistant to extreme values.
5. Understand the differences between mean and median in their interpretation and properties.
6. Understand how to select appropriate measures of center to represent a sample of data.

Description of the Lesson

In the first activity described at the beginning of the chapter (What does the Mean Mean Activity) students are told that the average age (mean) for students in the class is 21 years and are asked to think about what we know about the distribution of students’ ages for this class (e.g., “Are they all about 21 years old?”), and to explain their answer first in a small group and then to the class. They are also asked to explain where this value of 21 came from and how it was produced. Students are asked to make conjectures about the
ages of these 10 students and to use 10 Post-It® notes and to form a series of dot plots on
a given number line so that the average is 21 years.

Students move one Post-It® note to 24 years, and later one to 17, and figure out how to
move one or more of the other Post-It® notes to keep the mean at 21 years, discussing
their strategies and reasoning with their group and then the class. The term deviation is
introduced to represent the distance of each data value from the mean and students
examine deviations for their plots under different conditions, seeing how they have to
balance to zero.

In the second activity (What does the Median Mean?), students reduce their Post-It®
notes to 9 and arrange them on the same number line used earlier so that the median is 21
years. Again, they are given different constraints (e.g., change one of the values that is
currently higher than 21 years) and determine if and how the median affected. Finally,
students are asked to discuss and summarize what would they have to do with a data
value in the plot in order to change the median.

In the final activity of this lesson (Means and Medians), students observe and discuss two
Fathom Demos: The Meaning of Mean and Mean and Median (adapted from Fifty
Fathoms, Erickson, 2002) to further understand properties of these measures. The lesson
ends with a wrap up discussion about use interpretation, and properties of the mean and
median,

Lesson 2: Choosing Appropriate Measures

This lesson introduces the idea of choosing an appropriate measure of center to describe a
distribution. It has students predict typical values for variables that have different
distributions. The lesson then has students find the actual mean and median for those
variables using computer software and examine the distributional features that made their
prediction closer to either the mean or median. It also introduces the idea of outlier
influence on these measures of center. Student learning goals for this lesson include:

1. Deeper conceptual understanding of mean and median.
2. Understand when it is better to use each as a summary measure for a distribution of data.

3. How to generate these statistics using *Fathom* Software.

*Description of the Lesson*

Students are first asked how we can describe the typical college students taking an introductory statistics course, and in what ways do students in this class differ? They are asked to think about and discuss how people use the words: typical, average, and normal in an everyday sense and how these words are used as statistical terms: mean, median, center, and average.

In the *What is Typical Activity*, students are asked to consider a set of variables that were measured on their first day of class Student Survey. Working in pairs they are asked to make a prediction about what they might expect as a *typical value* for all students enrolled in their statistics class this term. They are reminded that a typical value is a single number that summarizes the class data for each variable. They write their prediction in the “First Prediction” column of the table shown below (Table 2).

*Table 2: Predicting and verifying typical values in the What is Typical Activity.*

<table>
<thead>
<tr>
<th>Attribute from Student Survey</th>
<th>First Prediction</th>
<th>Revised Prediction</th>
<th>Statistics from Fathom</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of statistics courses you are taking this semester</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Credits registered for this semester</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Next, they generate dot plots of the data using *Fathom* software to see if their original predictions seemed reasonable. Based on the graphs, they are allowed to make revised predictions for the typical value for each of the variables which are written in the table above in the “Revised Prediction” column.

Students next use *Fathom* to find the mean and median for each of these variables and complete the last two columns of the table above. They are asked to discuss how close were their revised predictions to the “typical” values produced by *Fathom* and for which attributes were their predictions most accurate. They are also asked what results were most surprising to them and why, and whether in general, were their revised predictions closer to the means or medians of these variables.

Students are asked to consider:

- Which measures of center were closest to their intuitive ideas of “typical” values?
- What information do means and medians provide about a distribution?
- How to decide whether to use the mean or median to summarize a data set?
- In statistics, what is meant by the word “typical”? 

<table>
<thead>
<tr>
<th>Total college credits completed</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Cumulative GPA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hours a week you study</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of emails you send each day</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of emails you receive each day</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
In *Choosing an Appropriate Measure of Center Activity*, students are asked to suggest conditions where the mean and median provide similar information and when they give different information for the same data set. This leads to a discussion of which measure is more appropriate for each variable and why, and how to choose the best measure of center for a data set.

Students are asked if it is all right to compute a mean or median without first looking at a graph of data and then why that is not a good idea. They are asked to reason about what information is missing if all they are given is a measure of center, including what they know and not know if all they were given were measures of center. This provides a segue to discussion on spread (the next unit) and reinforces the connection between center and spread. In a wrap-up discussion students are asked to imagine a variable that could be measured in two different settings that might yield data sets that have the same mean and different amounts of spread, one with a little spread and one with a lot of spread, and explain their reasoning.

**Summary**

The two lessons in the unit on measures of center are closely connected to ideas of distribution and variability, so that the ideas of mean and median are always connected to these concepts and contexts. The intent of the lessons is to help students build a conceptual understanding of mean and median as well as the idea of center of a distribution, through physical manipulations of data values, making and testing conjectures about typical values, and discussing the use and properties of these two measures. While the concepts may seem simple, and not worth two full lessons, we believe that these lessons provide important foundations for and connections to subsequent units in the course.

**References**


Chapter 10: Learning to Reason about Variability

Variation is the reason why people have had to develop sophisticated statistical methods to filter out any messages from the surrounding noise. (Wild & Pfannkuch, 1999, pp. 235 - 236)

Snapshot of a Research-based Activity on Variability

Students are given pairs of histograms and are asked to discuss, which graph in each pair would have a higher standard deviation and why. First students are encouraged to actually draw deviations on the histograms and actually draw lines from the mean to represent the number of deviations for each bar of the histogram (e.g., a bar representing five values would correspond to five lines of the same length, showing deviations from the center). Some of the histograms are easier to compare than others, such as those that have a few bars close the center versus a histogram with most bars far from the center (see examples in Figure 7 below). The difficult comparisons are for graphs that have the same range, the same frequencies for each bar, but different numbers of bars (representing different possible values of the variables).

After students compare and discuss their answers, they enter the data for each graph into Fathom (Key Curriculum Press, 2006) and have the standard deviations computed. They use Fathom to show the actual squared deviations from the mean for each graph. After the students have checked their answers in this way, the class discusses each pair of histograms and why the standard deviations were large or smaller in each pair. As they do this, they construct a set of factors that appear to influence the size of the standard deviation (e.g., more bars farther from the mean) and those that do not seem to affect the size (e.g., “bumpiness” of the graph, or different heights of the bars).

Rationale for This Activity

While students learn what the standard deviation is and how it is calculated, they rarely have an understanding of what this measure is and how to interpret it. The activity
described above is a culminating activity in the unit on variability, because it helps students recognize and integrate several sub ideas (e.g., a graphical representation of distribution, the mean of a distribution, spread, and deviation). The idea of standard deviation as an average distance from the center is developed by first having students estimate the mean of each graph, taking into account both value (on the number line) and density (frequency of each bar). Students are guided to reason about deviations from the mean and think about how values close to and far from the mean affect those deviations and squared deviations.

Entering the data from each histogram themselves (to find the actual standard deviations for each pair of graphs), helps students remember that each bar in the histogram represents one or more pieces of data of the same value, distinguishing these graphs from case value graphs (see more on this issue in Chapter 8). Seeing the squared deviations generated by Fathom illustrates how deviations far from the mean have a greater impact on the size of the standard deviation.

The Importance of Understanding Variability

*Variability is ... the essence of statistics as a discipline and it is not best understood by lecture. It must be experienced* (Cobb, 1992).

Any introductory course should take as its main goal helping students to learn the basics of statistical thinking (Cobb, 1992) which includes the omnipresence of variability and the quantification and explanation of variability. These two topics are highlighted in the GAISE report (2005):

*The omnipresence of variability*

Recognizing that variability is ubiquitous. It is the essence of statistics as a discipline and it is not best understood by lecture. It must be experienced.

*The quantification and explanation of variability*
Recognizing that variability can be measured and explained, taking into consideration the following: (a) Randomness and distributions; (b) patterns and deviations (fit and residual); (c) mathematical models for patterns; (d) model-data dialogue (diagnostics).

Understanding the ideas of spread or variability of data is a key component of understanding the concept of distribution, and is essential for making statistical inferences. While students develop informal ideas of spread in the earlier unit on graphing and describing distributions, they later encounter these ideas more formally as they learn about different measure of variability (e.g., range, standard deviation and interquartile range), what they mean, how to interpret them, how they compare to each other as statistical summaries of data, and what information they provide and don’t provide, and how we use them in analyzing data.

There has been increasing attention paid to the importance of students developing an understanding of and appreciation for variability, as a core component of statistical thinking (Cobb, 1992; Moore, 1998). However, it is impossible to consider variability without also considering center, as both ideas are needed to find meaning in analyzing data.

**The Place of Variability in the Curriculum**

The idea of spread, or variability should permeate the entire curriculum. We advocate introducing ideas of spread first informally, and later formally. Ideas of variability can be introduced the first day of class (see lessons from the data unit) and revisited in the unit on distribution, where students describe the spread or clustering of values in a graph of a distribution. When center is introduced, the idea of deviation from the mean is used to help understand the meaning of the mean, and this idea of deviation form the mean is then revisited when studying standard deviation. While range is a fairly easy concept for students to understand, standard deviation is much more difficult. Interquartile range is also a difficult concept, and best introduced in the context of comparing groups with boxplots, when it is illustrated visually by the width of the box.
It is hard to imagine a situation where one would summarize a data set using only a measure of center or using only a measure of spread. When comparing groups or making inferences we need to look at center and spread together: the signal, and the noise around the signal. Therefore, ideas of center and spread are most often seen and used together, whether informally describing distributions, looking at theoretical models such as the normal distribution and sampling distributions, or in making inferences.

**Review of the Literature Related to Reasoning about Variability**

*Variation vs. Variability*

Before we begin to summarize current research on reasoning about variability, we want to address the question of terminology. An inconsistent use of terminology is noticeable in research studies about variability. While some use “variability” and “variation” interchangeably, others distinguish between the meanings of these two words. Reading and Shaughnessy (2004) suggest the following distinctions: variation is a noun describing the act of varying, while variability is a noun form of the adjective “variable,” meaning that something is likely to vary or change. They suggest that variability refers to the characteristic of the entity that is observable, and that variation refers to the describing or measuring of that characteristic. We have chosen to use the term *variability* as the general, omnibus term for these ideas in this chapter.

*The Emergent Research about Variability*

Recent discussions in the statistics education community have drawn attention to the fact that statistics text books, instruction, public discourse, as well as research have been overemphasizing measures of center at the expense of variability (e.g., Shaughnessy, 1997). Instead, there is a growing consensus to emphasize general distributional features such as shape, center and spread and the connections among them in students’ early experiences with data. It is also suggested to focus students’ attention on the nature and sources of variability of data in different contexts, such as variability in a particular data set, outcomes of random experiments, and sampling (Shaughnessy, Watson, Moritz, & Reading, 1999; Gould, 2004). These views are supported by a review of several studies.
by Konold and Pollatsek (2002) that has shown that “the notion of an average understood as a central tendency is inseparable from the notion of spread” (p. 263). Their well-known metaphor for data as signal and noise implies that students should come to see statistics as “the study of noisy processes – processes that have a signature, or signal” (p. 260).

Difficulties in Understanding Variability

Despite the widespread belief in the importance of this concept, current research demonstrates that it is extremely difficult for students to reason about variability and that we are just beginning to learn how reasoning about variability develops (Garfield & Ben-Zvi, 2005). Understanding variability has both informal and formal aspects, moving from understanding that data vary (e.g., differences in data values) to understanding and interpreting formal measures of variability (e.g., range, interquartile range, and standard deviation). While students can learn how to compute formal measures of variability, they rarely understand what these summary statistics represent, either numerically or graphically, and do not understand their importance and connection to other statistical concepts. What makes the understanding of the concept even more complex is that variability may sometimes be desired and of interest, and sometimes be considered error or noise (Gould, 2004; Konold & Pollatsek, 2002), as well as the interconnectedness of variability to concepts of distribution, center, sampling, and inference (Cobb, McClain, & Gravemeijer, 2003).

These difficulties are evident, for example, in a series of interview studies with undergraduate students who had earned a grade of A in their college statistics course, Mathew and Clark (2003) found that students could not remember much at all about the standard deviation. In another interview study of introductory statistics students’ conceptual understanding of the standard deviation, delMas and Liu (2005) designed a computer environment to promote students’ ability to coordinate characteristics of variation of values about the mean with the size of the standard deviation as a measure of that variation. delMas and Liu found that students moved from simple, one-dimensional understandings of the standard deviation that did not consider variation about the mean to
more mean-centered conceptualizations that coordinated the effects of frequency (density) and deviation from the mean.

In a study investing students statistical reasoning, using the statistical reasoning Assessment (SRA), Garfield (2003) found that even students in introductory classes that were using reform textbooks, good activities, and high quality technology, had significant difficulty reasoning about different aspects of variability, such as representing variability in graphs, comparing groups, and comparing the degree of variability across groups.

**Developing Students Reasoning about Variability**

A variety of contexts have been used in statistics education to study students’ reasoning about variability at all age levels. For example, in a study of elementary students, Lehrer and Schuble (in press) contrast students’ reasoning about variability in two contrasting contexts: (a) measurement and (b) “natural” (biological). When fourth-graders were engaged in measuring the heights of a variety of objects, distribution emerged as a coordination of their activity. They were able to invent statistics as indicators of stability (e.g., center corresponded to “real” length) and variation of measure (e.g., spread corresponded to sources of error such as tool, person, trial-to-trial). In the context of natural variation activity (growth of plants), these same students (now fifth-graders) had difficulties handling sources of natural variation and related statistics. Activities that promoted investigations of sampling (e.g., “what would be likely to happen to the distribution of plant heights if we grew them again”), and comparing distributions (e.g., “how one might know whether two different distributions of height measurements could be considered ‘really’ different”) were found useful in developing students understanding of variability.

In a design research conducted with students in grades 7 and 8 (Bakker, 2004), instructional activities that could support coherent reasoning about key concepts such as variability, sampling, data, and distribution were developed. Two instructional activities were found to enable a conceptual growth: A “growing a sample” activity that had students think about what happens to the graph when bigger samples are taken, and an activity requiring reasoning about shape of data.
The advantage in discussing ideas of variability in connection with ideas of center was described by Garfield, delMas, & Chance (in press). In this study with undergraduate students, results indicated that students could develop ideas of a lot or a little variability when asked to make and test conjectures about a series of variables measuring minutes per day spent on various activities (e.g., time spent studying, talking on the phone, eating, etc.). They also found that by having students reason about the distributions of these variables informally they could move them for comparisons of formal measures of variability (e.g., standard deviation, range and interquartile range).

Other contexts examined include variability in data (Ben-Zvi, 2004a; Groth, 2005; Konold & Pollatsek, 2002; Petrosino, Lehrer, & Schauble, 2003), bivariate relationships (Cobb, McClain, & Gravemeijer, 2003; Hammerman & Rubin, 2003), comparing groups (Ben-Zvi, 2004b; Biehler, 2001; Lehrer & Schauble, 2002; Makar & Confrey, 2005), probability contexts (Reading & Shaughnessy, 2004), measures of spread such as the standard deviation (delMas & Liu, 2005), and sampling (Chance, delMas & Garfield, 2004; Watson, 2004). These studies are mostly exploratory and qualitative, and their research goal is often to explore what and how students come to understand ideas of variability in the different contexts. The kinds of questions and activities used in these studies suggest ways we can help students develop reasoning about variability across an entire course as well as assess informal and formal aspects of students understanding of variability.

It appears that it is difficult to help students move from informal to formal notions of variability without carefully designed activities leading them through this process, and have difficulty seeing variability as a fundamental idea underlying statistics. It is also difficult for students to recognize the different “faces” of variability, such as overall spread, clustering to the center, or relative deviations from an expectation (e.g., a measure of center or an expected model).

**Implications of the Research: Teaching Students to Reason about Variability**
Noticeably lacking in the current research literature are studies of how to best impact the learning of college students who are typically introduced to variability in a unit of descriptive statistics, following units of graphing univariate data and measures of center. Measures of variability (or spread) are then introduced, and students learn to calculate and briefly interpret them. Typically, only the formal notion of variability as measured by three different statistics (i.e., the range, interquartile range, and standard deviation) is taught. Students often do not hear the term “variability” stressed again until a unit on sampling, where they are to recognize that the variability of sample means decreases as sample size increases. When students are introduced to statistical inference, variability is then treated as a nuisance parameter because estimating the mean becomes the problem of importance (Gould, 2004).

Given this typical introduction in textbooks and class discussion, it is not surprising that few students actually develop an understanding of this important concept. Good activities and software tools designed to promote an understanding of variability do exist. However, they are typically added to a lesson or given as an assignment instead of being integrated into a plan of teaching, and their impact on student understanding has not been subjected to systematic study. So while there have been positive changes in introductory statistics classes, they still fall short of giving students the experiences they need to develop statistical thinking and a deep understanding of key statistical concepts.

We would like our students to follow the way statisticians think about variability. When statisticians look at one or more data sets, they often appraise and compare the variability informally and then formally, looking at appropriate graphs and descriptive measures. They look at both the center of a distribution as well as the spread from the center, often referring to more than one representation of the data to lead to better interpretations. Statisticians are also likely to consider sources of variability, including the statistical and measurement processes by which the data were collected.

Konold and Pollatsek (2002) offer the following suggestions about how we might help students and future teachers develop ideas of the signal-noise perspective of various statistical measures:
1. Using processes involving repeated measures;

2. Explorations of stability such as drawing multiple samples from a known population and evaluating particular features, such as the mean, across these replications.

3. Comparing the relative accuracy of different measurement methods;

4. Growing samples - students observe a distribution as the sample gets larger;

5. Simulating processes – students investigate why many noisy processes tend to produce mound-shaped distributions;

6. Comparing groups; or

7. Conducting experiments.

Garfield and Ben-Zvi (2005) outline a list of increasingly sophisticated ideas for constructing "deep understanding" of variability. This list offers the ways and order in which this body of knowledge can be structured and revisited as students progress through the statistics curriculum.

1. Developing intuitive ideas of variability

2. Describing and representing variability

3. Using variability to make comparisons

4. Recognizing variability in special types of distributions

5. Identifying patterns of variability in fitting models

6. Using variability to predict random samples or outcomes

7. Considering variability as part of statistical thinking
Rather than present material in a logical fashion, as most textbooks and course do, such an approach to teaching introductory college statistics might help students develop a deep understanding of the concepts of center and variability, understand how concepts are connected, and build their statistical thinking.

**Progression of Ideas: Connecting Research to Teaching**

*Introduction to the Sequence of Activities to Develop Reasoning about Variability*

The following table shows a series of steps that can be used to help students first build informal and then formal ideas of variability. These ideas are first introduced in earlier units on data, distribution and center. Then the formal idea of standard deviation is introduced and used to examine and reason about data. The concept of interquartile range is introduced in the later unit on comparing groups, a unit that helps connect ideas of center and spread visually and for the purpose of comparing sets of data to answer a research question. The idea of variability is visited again in the unit on models, when the normal distribution is introduced and the unique characteristics of the mean and standard deviation are shown as part of the Empirical Rule. The interconnections of center and spread are also demonstrated in the sampling, statistical inference and covariation units. Each time the basic idea of variability is explicitly revisited in that particular context, emphasized and discussed.

*Table 1: Sequence of activities to develop reasoning about variability.*

<table>
<thead>
<tr>
<th>Milestones: Ideas and Concepts</th>
<th>Suggested Activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>INFOROMAL IDEAS PRIOR TO FORMAL STUDY OF VARIABILITY</td>
<td></td>
</tr>
<tr>
<td>• Data vary. Values of a variable illustrate variability.</td>
<td>• Meet and Greet Activity (Lesson 1, Data Unit, Chapter 6)</td>
</tr>
<tr>
<td>• Variability in results from a random experiment.</td>
<td>• Activities in Lessons 1 and 2, Statistical Models Unit (Chapter 7)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Informal idea of spread of data by examining a graph.</th>
<th>Distinguishing Distributions Activity (Lesson 1, Distributions Unit, Chapter 8).</th>
</tr>
</thead>
<tbody>
<tr>
<td>Range as a simple measure of spread.</td>
<td>An activity where students describe distribution and note range as a measure of spread. (The symbol ❖ indicates that this activity is not included in these lessons.)</td>
</tr>
</tbody>
</table>

**FORMAL IDEAS OF VARIABILITY**

<table>
<thead>
<tr>
<th>Two ideas of variability: diversity or measurement error.</th>
<th>How Big is Your Head Activity (Lesson 1: “Variation”)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sources of variability, a lot and a little variability.</td>
<td>How Big is Your Head Activity (Lesson 1)</td>
</tr>
<tr>
<td>Averaging deviations from the mean as a measure of spread.</td>
<td>Comparing Hand Spans Activity (Lesson 2: “Reasoning about the Standard Deviation”)</td>
</tr>
<tr>
<td>Standard deviation as a measure of average distance from the mean.</td>
<td>Comparing Hand Spans Activity (Lesson 2)</td>
</tr>
<tr>
<td>Understanding factors that cause the standard deviation to be larger or smaller.</td>
<td>What Makes the Standard Deviation Larger or Smaller Activity (Lesson 2)</td>
</tr>
<tr>
<td>How center and spread are represented in graphs?</td>
<td>An activity where students match a set of graphs to the corresponding set of statistics.</td>
</tr>
</tbody>
</table>
Building on Formal Ideas of Variability in Subsequent Topics

<table>
<thead>
<tr>
<th>• Range and IQR in a boxplot.</th>
<th>• How Many Raisins in a Box Activity (Lesson 1, Comparing Groups Unit, Chapter 11)</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Variability within a group and variability between groups.</td>
<td>• Gummy Bears Activity (Lesson 2, Comparing Groups Unit, Chapter 11)</td>
</tr>
<tr>
<td>• What makes the range and IQR larger and smaller?</td>
<td>• How do Students Spend their Time Activity (Lesson 4, Comparing Groups Unit, Chapter 11)</td>
</tr>
<tr>
<td>• Understanding how and why center and spread are used to compare groups.</td>
<td>• How do Students Spend their Time Activity (Lesson 4, Comparing Groups Unit, Chapter 11)</td>
</tr>
<tr>
<td>• Role of mean and standard deviation in describing location of values in a normal distribution.</td>
<td>• Activities in Lesson 3, Statistical Models Unit (Chapter 7)</td>
</tr>
<tr>
<td>• Understanding why and how variability decreases as sample size increases in sampling distributions.</td>
<td>• The Central Limit Theorem Activity, (Lesson 3, Samples and Sampling Unit, Chapter 12)</td>
</tr>
<tr>
<td>• Understanding ideas of variability between and within groups when comparing samples of data.</td>
<td>• Gummy Bears Revisited Activity (Lesson 4, Statistical Inference Unit, Chapter 13)</td>
</tr>
<tr>
<td>• Variability of data in a bivariate plot.</td>
<td>• Interpreting Scatterplots Activity (lesson 1, Covariation Unit, Chapter 14)</td>
</tr>
</tbody>
</table>

Introduction to the Lessons
While students have been informally introduced to the idea of spread and range earlier, this set of lessons looks more closely at variability and the standard deviation. First students collect measurements to help them recognize two ways of looking at variability, as noise and as diversity. They informally think about a measure of spread from the center. The second lesson helps develop the idea of standard deviation and encourages students to reason about how this statistics is used to measure and represent variability and factors that affect the standard deviation, making it larger or smaller.

**Lesson 1: Variation**

This lesson is designed to help students reason informally about variability. Students compare measurements for two sets of repeated measurements, to discover two kind of variability: 1) variability as an error of measurement (repeated measures of the same head circumference); and 2) variability as an indicator of diversity (measurements of different people’s head circumferences). Students are then introduced to the concept of signal and noise, and discuss the stability of the mean as more data are collected. Student learning goals for the lesson include:

1. Understand different types (sources) of variability (when it’s desired and when it’s noise).

2. Understand the ideas of mean as signal and variability as noise, from repeated measurements in an experiment.

3. Understand that it is desirable to reduce variability in measurement (by using experimental protocols).

**Description of the Lesson**

Students are asked to think about variability in the class, and in particular, of head sizes. In the *How Big is Your Head Activity*, they plan a method to measure the circumference of each of their heads, keeping track of the decisions they make about measuring. A class discussion about this results in a common protocol to use. Students are given a measuring
tape to use, and they measure each of their heads using the protocol established and they record the data on the board (or on a computer spreadsheet).

Next, as a class, the students choose one person who will have their head measured by every student in the class. These measurements are also recorded for the class. Students then work with a partner to obtain a set of additional body measurements (all in centimeters) listed in Figure 1. These data will be used in other activities in the course. These data are later entered in Fathom.

**Body Data Collection**

- **Height** (with shoes on): __________
- **Arm Span** (from fingertip to fingertip with arms out-stretched): __________
- **Kneeling Height**: __________
- **Hand Length** (from the wrist to the tip of the middle finger): __________
- **Hand Span** (from the tip of the thumb to the tip of the pinkie while hand is stretched): __________

*Figure 1: Record sheet for the body measurements survey.*

*Fathom* is used to create a plot of students’ head sizes so that they may be examined and summarized. Unusual values are examined and discussed to see if they are legitimate or the result of a faulty measurement process.

Students are then asked to select two numbers that seem reasonable for completing the following sentence. (Note: There is more than one reasonable set of choices.)

*The typical head circumference for students in this class is about ______ cm give or take about ______ cm.*

These answers as discussed and lead to a discussion of possible reasons for the variability in the measurements of students’ head circumferences. Students are then asked to think
about whether the observed variability could be reduced and if so, how that might happen. They offer suggestions for ways to make the measurements more standard.

Next, the class examines data for the repeated measurements of one student’s head circumference. These data are entered into Fathom and are graphed and summarized in terms of shape, center and spread. This graph is then compared to the first graph of all students head circumferences and reasons for the differences are discussed. This time, the students suggest that the variability is solely due to the measurement process and talk about ways to reduce that variability.

A class discussion of the difference between these two sets of measurements of head circumference includes the different types and sources of variability, and when we might expect (and accept) variability in measurements and when we want to keep it as small as possible. The concepts of signal and noise are revisited, and the idea of variability as noise in the case of the repeated measurements of one head is discussed.

A wrap up discussion includes suggestions for different sources of variability in data, two kinds of variability: “diversity” (spice of life) and “error or noise”. Students are asked which type we like to have large and which do we like to have small, and why. They are finally asked to come up with some other examples of signal and noise, and to consider what we care about signal and noise when we examine data.

**Lesson 2: Reasoning about the Standard Deviation**

This lesson encourages students to reason about the standard deviation. Students begin by visualizing and estimating average distances from the mean without introducing the mathematical formula. First this is done in an activity involving hand spans. The lesson also has students compare graphs of distributions and reason about which of two graphs would have more variability (higher standard deviation). This lesson is designed to help students improve their reasoning about and understanding of variability by thinking about what a standard deviation is and applying that knowledge to determine which of two graphs has a higher standard deviation. Student learning goals for the lesson include:
1. Understand and informally estimate deviations from the mean and “typical” deviation from the mean.

2. Understand standard deviation as a measure of spread

3. Understand what makes standard deviations larger or smaller, what types of graphs reveal different amounts of variation.

4. Reason about connections between measure of center and spread, and how they are revealed in graphical representations of data.

Description of the Lesson

In the first activity, *Comparing Hand Spans*, students are asked to examine and compare their hands and think about variability in hand spans. Students find the hand span for every person in their group (Figure 2). They use a dot plot to examine how these values vary. They are asked to suggest two sources of variability in these measurements, i.e., two reasons why the measurements are not all the same.

Next, students record initials above the dots to identify each case. They find the mean and mark it with a wedge (▲) below the correct place on the number line. They are asked to estimate how far each of their hand span measurements is from the mean of their group. They make a second dot plot, this time of the differences (deviations) from the mean for each student in their group, and find the mean of these differences (deviations). Using the idea of deviations from the mean, students are asked to suggest a “typical” distance (deviation) from the mean.

*Fathom* is used to re-create the dot plot and to check their calculations and to compute the standard deviation of the group’s hand span data. Students are asked to compare the actual standard deviation to the “typical” distance (deviation) the group found earlier and to speculate about the difference in these values.

Figure 2: Measuring hand span.
Students then access the entire set of Hand Spans for the class that were gathered in the pervious activity and find the standard deviation of these measurements. This statistic is compared to the standard deviation of hand spans for the small group of four originally produced, and differences are discussed. Finally, students discuss the idea of a “typical” deviation and the standard deviation.

The second activity What Makes the Standard Deviation Larger or Smaller is adapted from delMas (2001, STAR library, http://www.causeweb.org/repository/StarLibrary/activities/delmas2001/) and continues the discussion of a typical, or standard, deviation from the mean. First, students examine the following dot plot (Figure 3), which has the mean marked by a vertical line. Students are asked to think about how large the deviations would be for each data point (dot).

![Figure 3: Dot plot from the What Makes the Standard Deviation Larger or Smaller Activity.](image)

Next students draw in the plot each deviation from the mean as shown below (Figure 4).

![Figure 4: Drawing deviations from the mean in a dot plot from the What Makes the Standard Deviation Larger or Smaller Activity.](image)
Next, students are asked to reason about the average size (length) of all of those deviations, and use this to estimate the standard deviation. They draw the estimated length of the standard deviation. This process is repeated with a second dot plot as shown below (Figure 5).

![Second dot plot from the What Makes the Standard Deviation Larger or Smaller Activity.](image)

Figure 5: Second dot plot from the What Makes the Standard Deviation Larger or Smaller Activity.

Then students are given a histogram (Figure 6) and they are asked to use the same process, thinking about dots “hidden” by the bars, and to draw and estimate the length of the standard deviation. The mean of the data set is given (2.57). Students are encouraged to draw in the appropriate number of dots in each bar of the histogram to make sure they have the appropriate number of deviations.

![A histogram from the What Makes the Standard Deviation Larger or Smaller Activity.](image)

Figure 6: A histogram from the What Makes the Standard Deviation Larger or Smaller Activity.

Students are then given six pairs of histograms, for which they are to try to determine which graph in the pair would have a larger standard deviation or would they be the same, and why. The mean for each graph is given just above each histogram. In doing so,
students try to identify the characteristics of the graphs that make the standard deviation larger or smaller. Two such Pairs of histograms are shown below (Figure 7).

1. A \( \bar{x} = .33 \) B \( \bar{x} = 4.33 \)

   a. A has a larger standard deviation than B.
   b. B has a larger standard deviation than A.
   c. Both graphs have the same standard deviation.

2. A \( \bar{x} = 2.50 \) B \( \bar{x} = 2.56 \)

   a. A has a larger standard deviation than B.
   b. B has a larger standard deviation than A.
   c. Both graphs have the same standard deviation.

Figure 7: Comparing standard deviations of pairs of histograms from the What Makes the Standard Deviation Larger or Smaller Activity.

After students complete the set of comparisons, their answers can be discussed and compared as a class and correct answers provided (e.g., the actual size of the standard deviation for each graph in the activity). Students are asked to elaborate on which graphs were harder to compare, which were easier and why.

In a wrap-up discussion students are asked to comment on why we need measures of variability in addition to measures of center, and why variability is so important in data analysis. They speculate on why variability is the basis of statistical analysis and how we represent and summarize variability.
Summary

The two lessons in this unit focus mainly on the ideas of types of variability and the meaning of the standard deviation. If students can develop an understanding of this important measure of spread it will help them learn and reason about the related concept of sampling error in the units on sampling and margin of error in the unit on inference. The next unit (Chapter 11) introduces the range and interquartile range as measures of spread in the context of using boxplots to compare groups.

References


Biehler, R. (2001). “Girls (tend to) watch less television than boys” – Students’ hypotheses and data exploration strategies in group comparison tasks. Talk presented at LOGOS #10, Mathematics Education Unit, Department of Mathematics, The University of Auckland, New Zealand.


