Exploring Generalization

Unit

4

Mathematical Concepts

- Statistics invented to describe one sample distribution should work more generally to describe a wide range of sample distributions.
- A change in the process used to produce data generally results in a change in one or more characteristics of the resulting distribution.
- As particular characteristics (qualities) of a distribution change, so too should the statistics that describe those characteristics.
- In the case of using a better tool to measure the same attribute, the measure (statistic) of center should be relatively unaffected, but the measure (statistic) of precision should change.
- For production processes, measures of center estimate target values and measures of variability, the consistency of the process.

Unit Overview

This unit focuses on the important mathematical idea of generalization. Do statistics invented previously work well to describe new samples? When attributes of the distribution of sample values change, do the statistics also change in sensible ways? At the teacher's option, the attribute measured in Unit 1 (e.g., a teacher's arm-span) is measured again with another tool (e.g., a meter stick) that tends to produce less variation in the resulting collection of measurements. Students use statistics of center and variation developed in Units 2,3 to characterize this change, usually in variability (the precision of the measures) but not in center.

Generalization is promoted further by exploring production processes. A production process is one with a target value and with variation about the target value that arises just by chance. In manufacturing, products that are consistently produced to match target specification are valued. Statistics of center estimate the tendency of the production process to meet its target specification. Statistics of variability measure the consistency with which the target value is attained. Production processes featured in this unit are Toothpick Factory and Rate Walk. Teachers choose one or more production process for students to explore. Other production processes can be found in the Teacher's Corner of the modelingdata.org website.

Day 1: Taking New Measures
Students begin the unit by using a more precise measuring tool to measure the same length that they measured in Unit 1. This activity sets the context for considering how changes in process result in changes in distributions that can be described by statistics. Because the true measure of the

Contents

attribute does not change, the center is usually less affected than the variability of the resulting distribution of measures. If students have already measured the same attribute with a more precise tool and used these data in Units 2 and/or 3, then go to the section, Production Process.

Day 2: Comparing Distributions
Next, students make displays to compare the shape of distributions of measurements using more and less precise tools. Comparing displays helps viewers of the displays readily see what changes and what stays the same when the new tool is used.

Day 3: Trying Out Statistics
Students apply statistics developed in previous units to compare center and precision. Changes in the values of the statistics should reflect changes in the characteristics of the distribution.

Day 4: Production Processes: Toothpick Factory, Rate Walk
Toothpick Factory and Rate Walk help students think more generally about statistics of center and of variability. In these activities, statistics are used to make inferences about production processes. Production processes are used to create products, and statistics are used to measure the qualities of production, such as product consistency. In the Toothpick Factory, two different methods are used to package toothpicks. Students use statistics to estimate the target value per package and to decide which method is more consistent. In Rate Walk, students try to walk 10 meters at one of three target rates: ½ m/sec., 1 m/sec, and 5 m/sec. Students use statistics to compare their performances across the 3 different rates. How close do they come to target rates? Are they equally consistent across the three rates? How might they develop an indicator of which rate walk was best? The Rate Walk clarifies how accuracy (the target value) and precision are not identical. Rate walks can be consistently off target!

Days 4 and 5: Formative Assessment
Students respond to a 3-item quiz. Each item revisits conceptions of statistics. The focus is on differentiating measures of central tendency (mean, median) from measures of variability (range, inter-quartile range, average deviation) in contexts of repeated measure and production.

Materials & Preparation

Read

- ☐ **Conceptions of Statistics construct map**
  Read the construct map and/or visit the website (modelingdata.org) to view a progression of student thinking about statistics. Consider especially examples in Level 3, focusing on the forms of thinking and talk that demonstrate that students are beginning to understand statistics as replicable and generalizable (CoS3C-CoS3F).
- ☐ **Units 2 and 3**
  Reread Units 2 and 3, considering student thinking and teacher supports in relation to measures of center and precision. Although the statistics are the same, their interpretation is different for processes of production. Measures of center indicate an expected target value and measures of variability indicate the consistency with which the target value is attained.
- ☐ **Using TinkerPlots** Read about tools for investigating precision on the web site (modelingdata.org) and/or visit the TinkerPlots tutorials for the Measure, Divider and Hat Plot tools.
- ☐ **Read Discussion Guide and Statistical Generalization Conversation Planner** to plan a whole class conversation where students reinterpret statistics of center and variability as measures of target value and consistency of production for production processes.

Gather

- ☐ A new and more precise measuring tool, such as a yardstick or tape measure (see Taking New Measures). You will need only one, as students will take turns measuring.

If you choose to do the Toothpick Factory activity:

- ☐ For each pair or small group of students: Approximately 500 toothpicks in a bag, 9 plastic sandwich bags, 2 104 mm paper strips. and a bundle of exactly 50 toothpicks wrapped by a rubber band.

If you choose to do the Rate Walk activity:

- ☐ A timer
- ☐ A tape measure to lay out a 10m course
- ☐ Cones or some other markers for starting and finishing lines.

Materials & Preparation — Exploring Generalization Unit 4

Prepare

- ☐ **Decide how your class should begin this unit-measuring with a more precise tool or with a production process.**
  If students have not already measured the same attribute (e.g., arm span) twice, once with a relatively poor tool and once with a relatively good tool, then begin the lesson as suggested in Taking New Measures on the next page. If students have already taken measures using a more precise tool, consider which of the thought-revealing questions in the Taking New Measures lesson will be most useful for helping students begin to compare measures. Or, if you have already done this in Units 2, 3, then move on to Exploring Production Processes.

- ☐ **Consider how you would like students to create displays**.
  Decide whether you are going to have students create displays of the new measures by hand or use TinkerPlots. We recommend the use of TinkerPlots, because it is more efficient and still allows students to create and notate displays.

Taking New Measures

In this activity, students use a more precise measuring tool to measure again the attribute measured in Unit 1. They consider both what might make a measuring tool more precise (for instance, it does not have to be iterated as often, minimizing gaps and overlaps) and how measures of center and precision will be affected by the use of a more precise tool. The discussion in this activity provides another opportunity for students to relate qualities of a distribution to the process responsible for creating it. It is important that they make and explain their predictions about changes in center and precision before generating and comparing displays.

Note: If students have already used a more precise tool to measure the same attribute, generate data with a production process, such as the toothpick factory or the rate walk described on pages 17-19.

Whole Group

1. **Introduce the activity.**

   Remind students of the way they measured the length (arm-span or other body part) in Unit 1. Show students a new and more precise measuring tool—for example, a yardstick or a tape measure—that will minimize the number of times the tool is moved. Tell them they will measure the same attribute on the same person, but this time using the new tool.

   Note: You might select the new tool and show it to students, or you might want to engage them in suggesting a measurement tool that could increase the precision of the measurements of the attribute, as compared to those they collected in Unit 1.

2. **Involve students in considering how using a new tool might affect the measurements of the attribute.**

   Use thought-revealing questions to generate discussion. For example:

   Q: How do you think the measurement (of the attribute measured in Unit 1) can be improved?
   Q: (If students have already measured the same attribute twice) Which tool was better for measuring? Why?

Construct: MoV1(a) and MoV2(b)
This task engages students in thinking at the early levels of the Modeling Variability construct.

Q: If we use a better tool (hold it up), what will stay the same about our collection of measurements? What will change? Why do you think so?

Q: When we make displays of our new measurements, in what ways do you think the displays will look different? Can you show us by making a sketch? What makes you think the display will look different in this particular way?

Q: What do you think will happen to the best guess for the real measure of (the attribute measured in Unit 1)? Why do you think so?

Q: What do you think will happen to your precision number? Why do you think so?

Individual Work

3. Allow students to complete their measurements (if they need to do so).

Note: As in Unit 1, there are a variety of ways you can facilitate the collection of the measures so that measures are independent of each other and make efficient use of time.

4. Listen and watch for ideas to highlight in the whole-group discussion.

Note: You might hear students noticing that the measuring procedure is easier to carry out, or you might notice students are less likely to lose count or need to start over in their measures.

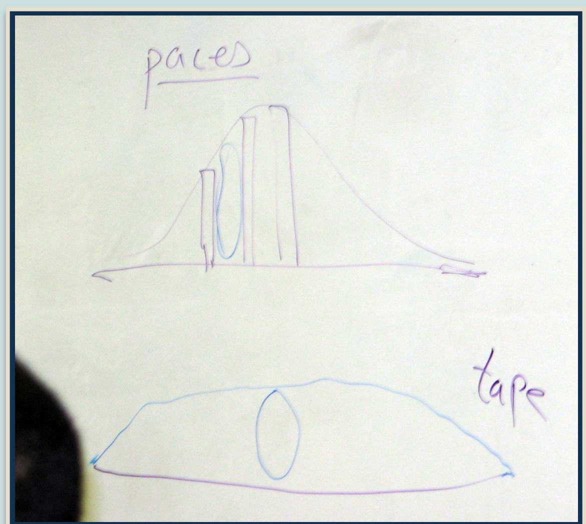
Whole Group

5. Close the activity by discussing any differences in measuring that students noticed.

Ask students to think about whether there are any reasons their measures would be more or less precise. This is a short discussion to provide more opportunities for students to relate their measures to the process by which the measure is created.

Students' Ways of Thinking: Predicting the results of a better measuring tool

When asked to think about the results of better measurement tools, students often predict, "everything will be closer." But it is a good practice to ask students to draw the shape of the data, because they occasionally think the shape of the data will be changed by the use of a different tool. For example, in one sixth-grade class, students drew the top panel in the photograph below to represent the results of the measurement of the height of the school's flagpole where distance on the ground was determined by footpaces. Students agreed this was a good approximation of the data obtained from their first measures with less precise tools. The bottom panel of the photograph represents student predictions about the effects of better tools (e.g., tape measures). The oval was drawn by a student to indicate denser, "closer" measurements in the "middle," but the shape of the distribution was also predicted to be more uniform. The instructor went on to explore with the class why they made this prediction.



A student prediction of shapes of data generated with different tools.

Comparing Distributions

In this activity, students make displays that allow them to compare the original collection of measures to the new collection of measures. It is important that they have displays of both collections to facilitate visual comparison and test their measures of center and precision. Beginning by qualitatively describing what they see in the displays helps to anchor their use of statistics (the best guess of the length of the attribute and the precision number) in the next activity to visual qualities of the displays.

Whole Group

1. **Introduce the activity.**

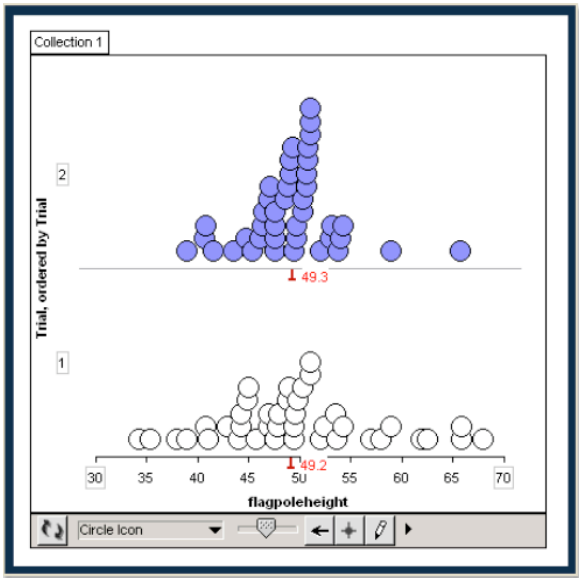
   Ask students to create displays of each sample of measurements, preferably using TinkerPlots, and to describe what they see.

   Note: It is useful to prepare a TinkerPlots file (either a table or collection of cards) that has both sets of measurements in it, differentiated by the name of the measuring tool (e.g., ruler and meterstick). It is OK if student names are not part of the file. But before sharing with students, challenge them to find a way to use TinkerPlots to enter the data in a way that will allow them to create a display that shows each measuring tool separately.

Collection 1

	Student	Tool	Measure...
1	Sara	Ruler	240
2	Sara	Meterstick	250
3	Jon	Ruler	270
4	Jon	Meterstick	248

By creating a separate attribute for the tool used, students can create displays that allow them to compare the two distributions easily, as in the figure on the next page, where data points are separated by measurement tool and positioned on a common scale (in the figure, Trial 1 refers to measuring with a ruler and Trial 2 refers to measuring with a meter stick).



Individual or Small Group Work

2. Have students make and describe their displays.

3. Observe and talk with students as they work.

Ask questions to understand what students are noticing about the displays and how they are describing center and precision.

Q: Tell me about how you are organizing the displays to compare the two collections of measurements. What does that allow you to see?
Q: When you compare the two collections, do you think they both show the same best guess of the length or not? What makes you think that?
Q: Is the new method of measuring more precise? Can you see evidence of that in the display? What makes you think that?

Whole Group

4. Describe and highlight differences in the two distributions.

Your goal is to help students relate similarities and differences in the center and precision of the two distributions to the processes involved in creating the measurements. You might ask questions such as these:

Q: Looking at your displays, were your predictions about what would stay the same and what would change confirmed?
Q: Did you see anything that you didn't expect? Why do you think that happened?
Q: Do you think the two displays show the same best guess for the length of the arm-span (or other measures)? What makes you think so?
Q: Do you think both methods for measuring are equally precise? What makes you think so?

Trying Out Statistics

In this activity, students use statistics they invented or used in Units 2 and 3 to compare the two methods for measuring. By this point in the unit, students should think that the new method (e.g., measuring with a meter stick) produced a similar best guess of the measure (mean or median) but is more precise; they should also have a sense that this is evident in the visual comparison of the two displays. They will now have the opportunity to see whether their statistics of center and precision reflect these qualities. In this activity, it is important for students to consider how the mathematical properties of their statistics show or hide the similarity in best guess and the difference in precision number of the two measurement methods.

Whole Group

1. **Introduce the activity.**

   Ask students to:

   a. Compare the best guess of the measure of the attribute for each distribution (the one generated using a crude tool and the one generated using a better tool).

   b. Compare the precision of measurement for each distribution.

   Note: You might choose to have students consider both center and precision measures on the same day or on different days. It is important that, across the class, several different measures of both center and precision are applied to the two sets of data. In particular, it is useful to encourage the use of range, interquartile range, and average or median deviation as statistics that measure precision.

Construct: CoS3(a) – CoS3(f)
These levels include examples of student thinking and, in the videos, teacher responses to those forms of thinking.

See Unit 3 for a description of measures of precision and teacher responses to student thinking.

Individual or Small Group Work

2. **Allow students to work alone or in partners.**

3. **Talk to individuals and groups and begin to think about which methods you would like to select for presentation and comparison.**

Note: As you move around the room, consider the different measures (statistics) that students are using and how each reflects (or fails to capture) the stability of the best guess and the difference in precision across the two sets of measurements. In particular, comparisons of methods of precision can lead to rich discussion. You might look for examples of the following:

Range-based measures. If some students employ a measure of the range (i.e., subtracting the lowest value from the highest value of the data set) their precision measures might or might not reflect the increased precision of the new method, depending on how the measurements are distributed. You might want to choose a range-based measure to share in the whole group discussion, if you have not yet discussed as a class why these measures can be problematic. You should ask students to consider that the range relies on the two most distinct measurements—and if one is concerned with tendency for measurements to agree, is that a good measure?

Measures that do and do not take into account changes in sample size. The two samples used here (more and less precise measurements) most likely have similar sample sizes. It is also important that methods generalize to and reflect differences in the characteristics of samples of different sizes. Pay attention to whether students' methods would allow them to compare the center or precision of samples of different sizes. Interquartile range, average deviation, or median deviation will allow for comparison of different sample sizes, while other methods may not.

Whole Group

1. **Use a whole group conversation to help students consider how different statistics generalize to the new sample and facilitate comparison of the two measurement methods.**

   See the description of Measure Reviews in Units 2 and 3 to consider how to plan for them and for instructional moves that are likely to be effective. Choose several individuals or groups to share their methods and report on how these methods allowed them to compare the center and/or precision of the two distributions.

   a. Tell students they will be focusing on a few of the methods they used. Tell them that while students are sharing their methods they

See the description of Measure Reviews in Units 2 and 3 to consider how to plan for them and for instructional moves that are likely to be effective.

should be thinking: Does this make sense? Does this help us to compare the two measurement methods? Could I do this on other data?

b. As each group presents, support them in making clear what method they used to calculate center and/or precision and to share how their methods reflected changes or stability in qualities of the two distributions. As needed, ask questions such as these:

Q: What happened to your best guess of the real measure? Did it change? By how much? Was this sensible? Why?
Q: Did your precision number change? How much? About how many times more precise were the measurements with the better tool compared to the measurements with the crude tool?

c. Allow students to ask each other questions. For example, it is often helpful for students to imagine changes in a distribution of data and ask what might happen to the value of a statistic.

d. Facilitate discussion of the measures by emphasizing aspects that make evident their mathematical properties. As needed, use questions like these:

Q: Which of the methods showed that the best guess of the measurement changed? Which did not? If some did and some did not, why is that?
Q: Which of the methods showed that the precision changed? Which did not? If some did and some did not, why is that?
Q: This method showed that the precision had changed by X, but this one showed that it changed by Y. Why is that? Is it a problem that the statistics show different changes?
Q: Which of the methods do you think are best for allowing us to compare between different measurement procedures? Why is that?

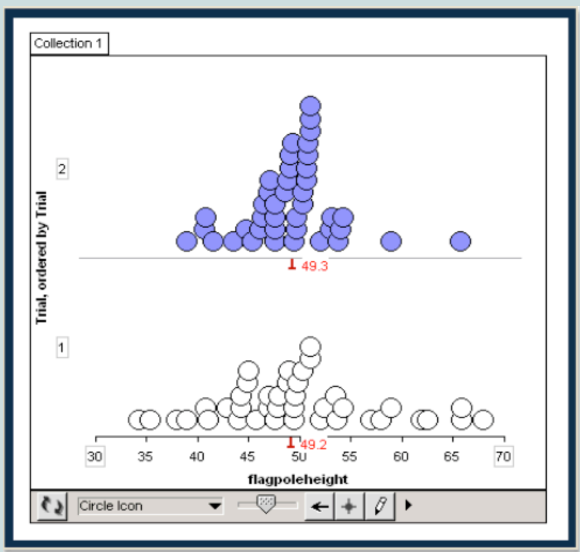
Note: Teachers have often found it useful to share “hypothetical distributions” to which the measures could be applied. Some that you might consider using are described below.

A distribution with a different range. If students employ range-based measures to compare the two measuring tools, an effective way to have students re-consider the generalization of the range is to ask what would happen if “one or two people really were careless or didn't know what they were doing." In this case, you might change the distribution of the more precise measurement method so that it had one or two outliers (values that are far from the others). Statistics that are based on range will portray the method as imprecise, though most of the values are still close together. Discuss with students why this type of measure is problematic in cases such as these. Compare it to interquartile range or average deviation measures.

A distribution with a different sample size. Changing the sample size wreaks havoc on measures that don't take it into account. You might prepare a tightly clumped set of data with double the sample size of your class's set of measurements, or a spread-out set of 6 measurements. Ask students to think about how the statistics they have presented would portray the precision of these sets of data. Be sure to discuss why particular methods don't facilitate comparison between samples of different sizes and whether the methods could be improved so that they do allow comparison of different samples.

Students' Ways of Thinking: Center and Precision

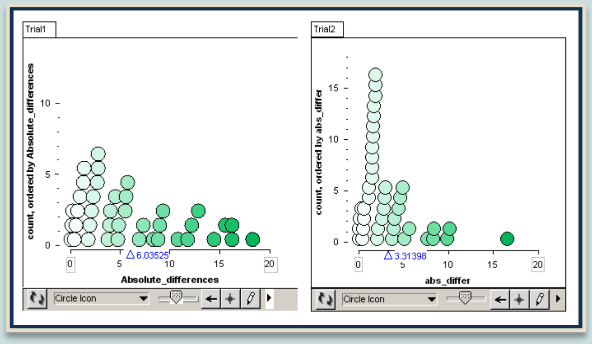
Comparing the central tendencies of distributions: Changing the measurement tool. Students usually notice that the central tendency of a collection of measures is not much affected by the change to a more precise tool. For example, in a sixth-grade class, the median value of the height of the flagpole using paces to estimate distance from the base of the flagpole was 49.2. The median value using a more precise measurement tool (i.e., a tape measure) was 49.3. The distributions are displayed below. Students didn't think this difference was sizable enough to matter. They suggested that the flagpole was not stretching or shrinking, so it made sense that the best guess of its measure was not changing.



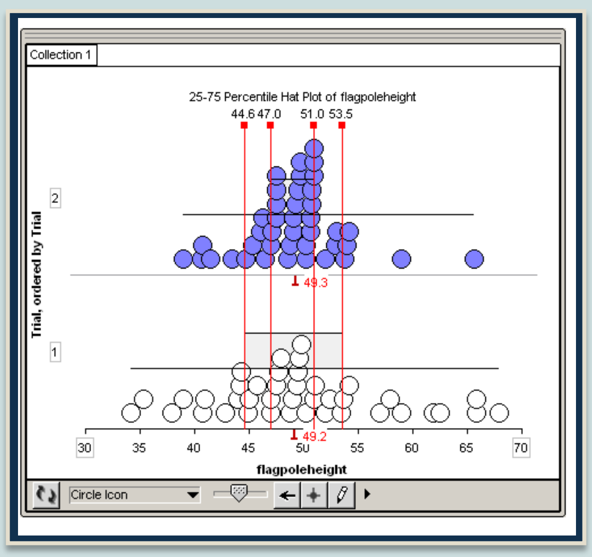
Two sets of measurements of flagpole height using two different tools

Comparing the Relative Precision of Two Collections of Measurements. Students always notice that using a better tool reduces the variability of the measurements. But the challenge is to determine whether or not invented measures of precision change to reflect this change in variability. Measures allow us to ask: How much better?

Attention to Distance from Center. A pair of sixth-grade students compared the relative precision of two collections of measurements by finding the average deviation of each one. The average deviation (the average of the absolute values of the differences between each observed value and the sample mean) of the first collection shown in the display below was 6.03 (left panel), and the average deviation of the re-measured collection was 3.31 (right panel). The instructor directed class attention to how the values of the statistic tracked changes in variability evident from visual inspection of the display.



Attention to Center Clump. Another student, who was reasoning about center clumps instead of deviation, created a display illustrated by the display below. Defining a center clump as the mid-50 percent of cases, he found the range of the center clump (the IQR), which decreased from pace measure (less precise tool) to tape measure. The IQR of the first trial with pace measure was 8.9 (53.5-44.6) and that of the second trial with tape measure was 4 (51-47). In this case, the lower IQR corresponds to more precise measurement.



A plot showing mid-ranges of the two trials

Production Processes

In this activity, students try out methods for finding center and precision in a new context of production processes. Production processes are critical in commerce where the uniformity of products is often taken for granted. People would not buy a car or even a candy bar that varied much from car to car or candy bar to candy bar. Statistics are used to help manufacturers control variability. In this section of the unit, two different but accessible production processes can be explored. Students investigate at least one production process. The first, Toothpick Factory, engages students in the production of packages of toothpicks. The packages are made in two different ways. Statistics of center estimate the target number for each package. Statistics of variability inform us about the consistency of each method of production. The second, Rate Walk, lets students know three different target rates for completing a 10 m walk. All students try to walk at each rate and the time elapsed for traveling 10 m is recorded. Statistics are used to compare rate walks: Were students more consistent at one of the rates compared to the others? Rate Walk clarifies the difference between accuracy and precision, because students can consistently walk at rates that are faster or slower than the target rate. For each production process, the teacher has the option of either generating data or using previously recorded data. We suggest, if time allows, generating data so that students have a good understanding of the processes used to generate the data. Whatever option you choose, you can use elements of the first three activities in this unit (facilitating the discussion of measures, creating and comparing displays, and trying out statistics) to guide your instruction, focusing on helping students consider how mathematical properties of statistical methods facilitate comparison and generalization.

Option 1: Toothpick Factory

See the Teacher's Corner on the modelingdata.org website for ideas about getting this activity started. Students work in partners or small groups to make packages of toothpicks. Each pair is given a kit containing 500 toothpicks, a bundle of 50 toothpicks wrapped by a rubber band, a strip of paper that is 104 mm long, 10 sandwich plastic bags, and a student sheet with directions about how to make the packages and a data log. Students use two different methods to package the toothpicks. The target number of toothpicks in each package is 50, though students are *not* told this. A student sheet ('Toothpick Factory') is included and can be found in the Student Worksheets section of this unit.

Taking New Measures
Comparing Distributions
Trying Out Statistics
Production Processes

The ***picket-fence method*** consists of placing the toothpicks side by side, vertically, and touching until the packing is 104 mm long (using paper strips cut to this length). Once the collection of toothpicks occupies the length of the paper strip, they are packaged: Put into a Ziploc bag and labeled to reflect the order of production (1, 2…10). The maximum time allowed is 6 minutes.

The ***bundle method*** consists of hefting the bundle of toothpicks and adding toothpicks to a bag until the weights feel the same. Each bag is labeled 1, 2, etc. The maximum time allowed is 6 minutes.

Each partner group uses both methods, or at the discretion of the teacher, different groups use different methods. After time expires, or while another member of the team is working, the number of toothpicks in each bag is counted and recorded. The toothpicks are then combined again to get ready for the next method. Encourage students to make sure each person has an opportunity to experience assembling the packages of toothpicks. Such experience will help them relate variability to process later in the activity.

When all data are collected, enter them into a Tinkerplots file, using the number of toothpicks counted in each bag, differentiated by which method was used (see Rate Walk, Option 2, for information on entering data into Tinkerplots so that it can be displayed and compared.)

Students can then try out their measures of center and precision to ascertain the target number of toothpicks per package and to consider which of the methods was more consistent. During whole-class discussion (see Discussion Guide on Modelingdata.org), start out with the combined data and ask questions such as these:

Q: Looking at all of our data, how many toothpicks should have been in each package? What statistics might be a good measure of this target number of toothpicks?
Q: Looking at all of our data, how consistent were the counts of toothpicks from package to package? Why weren't the counts of toothpicks exactly the same in all the packages? What statistics might be a good measure of consistency?

Then separate the data by method of production:

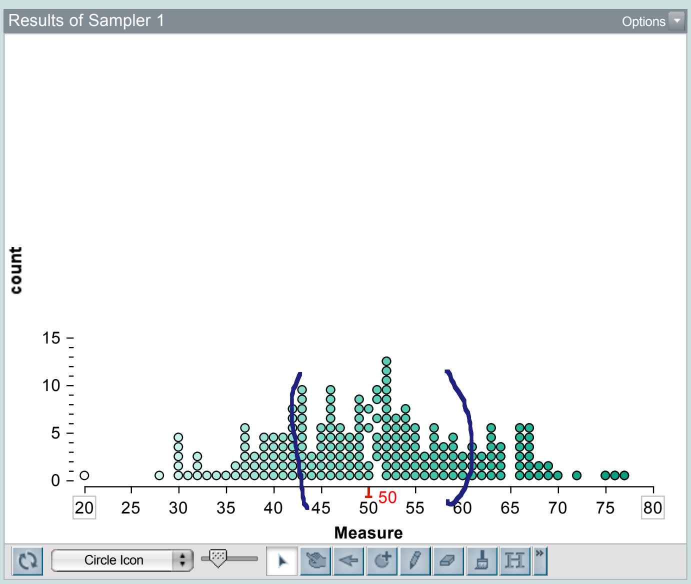
Q: Let's separate our data. How did the methods differ in terms of number of toothpicks per bag and the consistency of the number of toothpicks? What statistics tell you that?

Q: Which is the better packaging method? How do you know?

Q: What about the methods of production contributed to the differences in our results?

Students' Ways of Thinking: Center and Consistency

Target value and consistency. Students first decided that the median measured the target number of toothpicks per package. Students usually notice that one indicator of consistency of production is the "clumpiness" of the data around the target value. For example, a student below drew on a TinkerPlots display of all data obtained from the Toothpick factory to suggest her intuition about consistency.



After separating the data by method, and noticing that the target values did not vary very much, another student suggested that there might be a difference between methods in product consistency.

The class then turned its attention to developing a quantity that could represent this intuition—deciding upon the IQR. Another class selected the range as a good measure, but this measure was discarded after a student suggested that just a few "bad" packages in either method would produce about the same range, even though the packages were clearly more consistent with one of the methods, the picket fence.

Option 2: Rate Walk

Set up a course outside with a start line and a finish line separated by 10 meters. Students try to walk at a targeted rate for 10 meters and the time to complete the course is recorded. Students use the elapsed time to calculate a rate (10m / elapsed time). The target rates are 1 meter per second, 5 meters per second, and ½ meter per second.

Note: The activity is intended to elicit adaptation of statistics invented or taken up by students in the previous lesson. For example, a measure of closeness to each target rate might be the difference between the sample median time and the target time. A measure of consistency might be the inter-quartile range of the sample or the median deviation or the average deviation between each person's time and the target time. Decisions about which condition was best for meeting the target might use both measures. This is an opportunity for students to experience the distinction between accuracy in measurement, how close to the target, and consistency-precision, the tendency for the measurements to agree or cluster. If you would rather not have students try this activity for themselves, there is a student sheet in the Student Worksheets section of this unit ('Rate Walk') that provides a problem statement and table of data from a class of fifth-grade students who engaged in this activity. You can also download the data from the modelingdata.org website.

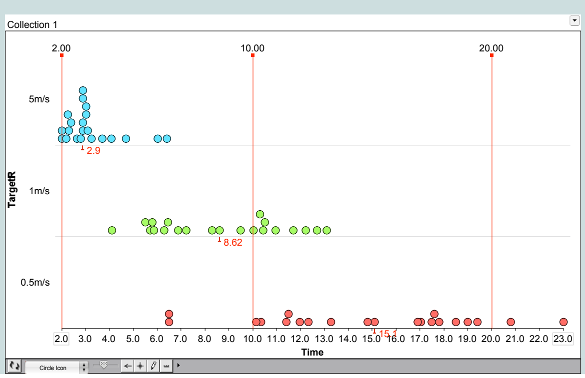
Students create distributions of the three conditions or create distributions the class can discuss together. They then apply their measures of precision and center to describe how close members of the class came to the target speed.

Useful questions to ask when teaching this activity include:

Q: How close do members of the class typically come to the target rate in each condition (1 m/s, 5 m/s, and ½ m/s)?
Q: How consistent are members of the class? Do the rates of the walks tend to agree?
Q: How can the answers to these questions help make a decision about which condition was best for meeting the target?

Students' Ways of Thinking: Target rates

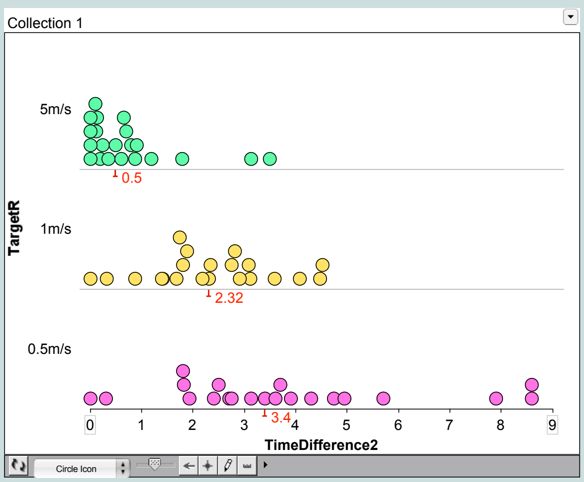
The image below shows a graph developed by a group of sixth-grade students to investigate rate walks. The students decided that the best estimate of how close they usually came to the target rate was to compare the median of the times to walk 10 meters for each condition to the target time (for example, 2 seconds at 5m/sec). They noticed that they came closest to the target at the fastest rate, likely because everyone ran just as fast as they could. The difference between the target and sample median was -0.9, indicating that they were usually nearly a second too slow. In the 1m/sec condition, the walkers tended to be too fast, with a difference of (10-8.6), 1.4 seconds too fast. The slow walk led to the worst approximation to the target rate, with a difference of (20-15.1) or nearly 5 seconds too fast. Walking slow was hard!



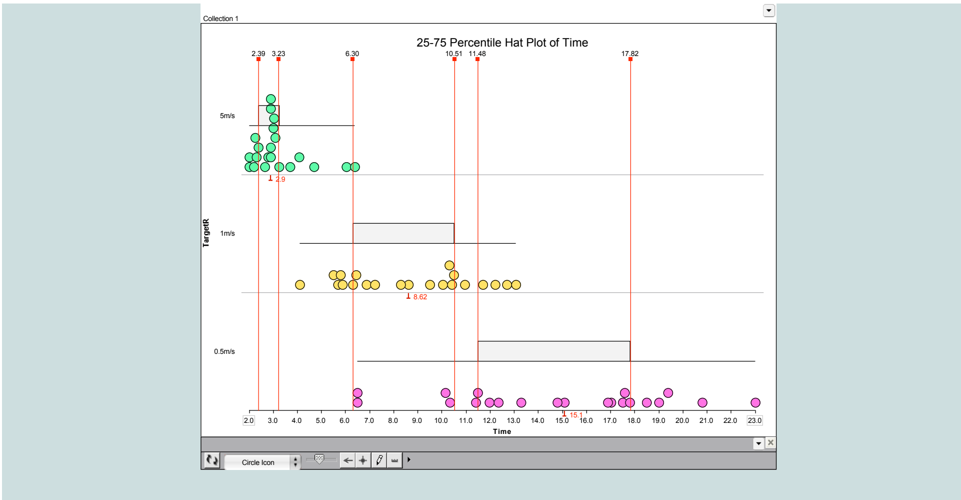
Taking New Measures
Comparing Distributions
Trying Out Statistics
Production Processes

When considering how consistently people walked, visual inspection again suggested that the tightest clumps were at 5m/sec (which should have taken 2 seconds to walk 10 meters). But to measure this tightness, students' inventions generally reflected two different ways of thinking about the clumps.

One method found the absolute value of the difference between every walker's time and the sample median in each condition and then found the average (the median) of these differences. The plot on the next page displays the collection of these differences and the median difference for each condition. Note that the median difference is least for 5m/sec.



A second method also used the logic of clustering about the median but instead found the mid-range (the interquartile range) of the middle 50 percent of the data. This method suggests that the tightest center clump is at 5m/sec (3.2 – 2.3 = 0.9), followed by the 1m/sec interquartile range of 4.2 (10.5 – 6.3), with the ½ m/sec again clearly the situation with the least consistent walks: 6.3 (17.8 – 11.5).


Collection 1
25-75 Percentile Hat Plot of Time
2.39
3.23
6.30
10.51
11.48
17.82
5m/s
1m/s
0.5m/s
TargetR
2.9
8.62
15.1
Time
Circle Icon

Formative Assessment

1. Administer the quiz.

Celo and Lexus measure the length of a parking lot with more precise and less precise tools. Students predict and justify whether the values of the mean and IQR will be about the same or be very different.

Speed Limit asks students to estimate a target speed limit from data and to measure the consistency with which drivers travel at this limit.

Acme Machines asks students to make an informal inference about whether or not a sample of data indicates that the Acme Company's production is consistent enough to merit keeping its contract. The statistic of interest here is the percent of the sample that does not meet the range of values specified by the manufacturer.

2. Use the scoring guides to score student responses.

3. Use *Celo's and Lexus'* measurements of the same length with different tools to link processes that generate a distribution with predictions about how these differences in tools will affect statistics of center and of variability (precision, consistency).

a. Select student responses to compare and contrast.

While scoring, select four different responses to use in the conversation, two related to the mean and two related to the average deviation. Be sure to include at least one student response that seems to suggest that mean and average deviation are both averages, so they are "the same," if this response is generated by any of your students.

b. Prepare questions to support and guide student thinking.

Q: What was being measured by each person?
Q: What is a mean? What does it measure here?
Q: What is the average deviation? What does it measure here?
Q: What might happen to a set of measurements if one tool is more precise than another?

c. **Use an Assessment Conversation to help students consider the relationship between a statistic and the process generating the distribution.**

Invite students to present their responses. Guide the conversation with questions that help the students understand that the mean estimates the real length of the parking lot. So, it should not change very much, if the way that each person measured was the same, with the exception of the tool that they could use. But the use of a different tool should affect the precision of measurement—the tendency of the measurements to agree. Better tools generate less error in measurement and therefore less variability. Relate this to students' experiences measuring a length with different tools.

4. **Use *Speed Limit* to generate a conversation about the appropriate statistics to use to estimate the target speed limit and to estimate the consistency with which motorists travel at this speed.**

a. **Select student responses to compare and contrast.**

Be sure to include responses that confound measures of center with measures of variability, if there are any. For example, some students may suggest the range as a good measure of the target value. Other students may calculate statistics, lumping together mean, median, mode and range, but may not be certain about which characteristic of the distribution each statistic measures.

b. **Prepare questions to support and guide student thinking.**

Q: In this situation, what is important to consider?
Q: If everyone followed the speed limit exactly, what would the display look like?

c. **Use an Assessment Conversation to help students consider the tradeoffs of methods.**
Questions that might stimulate further conversation might include:

Q: If we wanted to measure how consistently drivers tend to follow the speed limit, what is a:
(a) strength of the range? What might be a weakness?
(b) strength of the average deviation? What about the data does it pay attention to? What might be a weakness?

(c) strength of the interquartile range? What about the data does it pay attention to? What might be a weakness?

Q: For these data, which statistics are best to use? Why?

5. Use *Acme Machines* to generate a conversation about the appropriate statistic to use to make a decision.

a. Select student responses to compare and contrast.

Be sure to include responses that calculate statistics, such as medians or IQR's as well as those that find the percent of cases that are either less than 198 mm or more than 202 mm.

b. Prepare questions to support and guide student thinking.

Q: In this situation, what is important to consider?
Q: If the company made all the bolts exactly the same length, what would the graph look like?

c. Use an Assessment Conversation to help students consider the tradeoffs of methods.

Compare and contrast as students present their responses. Questions that might stimulate further conversation include:

Q: If there are 40 bolts made, how many of them can be more than 202 mm or less than 198 mm?
Q: What would a median value of 200 mm tell us? What doesn't it tell us?
Q: If we wanted to measure the consistency of the manufacture, what statistic could we use (IQR or average deviation)? For example, the IQR makes use of the tendency of the measurements.

The Toothpick Factory

You are working in a toothpick factory making packages of toothpicks. You are attempting to place the same number of toothpicks in each package. There are two methods that the factory is considering adopting. You will try out one of these methods in order to see how consistent it is and to discover the target value per package of toothpicks.

The **picket fence method** consists of placing the toothpicks side by side, vertically, and touching until the packing is 104 mm long (using the paper strip cut to this length). Once the collection of toothpicks occupies the length of the paper strip, they are packaged by putting them into a plastic sandwich bag. Label the bag in the order of production (1, 2…). The maximum time allowed is 6 minutes.

The ***bundle method*** consists of hefting the bundle of toothpicks held together by the rubber band and adding toothpicks to a bag until the weights feel the same. Label the bag in the order of production (1, 2…). The maximum time allowed is 6 minutes.

For the method you use, after you package the toothpicks, count and record the number of toothpicks in each bag.

Data Log

Bag #	Count	Method
1		
2		
3		
4		
5		
6		
7		
8		

Rate Walk

Students in one class tried to walk at different rates (1m/sec, 5 m/sec, ½ m/sec) without using any timers—just to see how close they could come. The class set up a course 10 meters long, and each person tried to walk at three different rates: 1 m/sec, .5 m/sec, and 5 m/sec. One person timed the walk. If a person walked exactly at 1m/sec, then it would take 10 seconds to travel 10m. If a person walked at 5m/sec, it would take 2 seconds, and at ½ m/sec, it would take 20 seconds.

Some people in the class said that people would be most accurate at a rate of 1 meter per second, but that anything slower (like ½ m per second) or faster (like 5 m per second) would be inaccurate. Other people said that accuracy would be about the same but that people would be most consistent (most like one another) at 1 m per second. It was hard to meet the target times, as you can see (The data are displayed in the table.)

Your job is to make a measure of (a) how close people tended to be to the target rate and (b) how consistent people were at each rate. Then, using your measures, make a decision about which rate seems easiest and which seems hardest.

STUDENT	1 meter per second	5 meters per second	½ meter per second
Makinda	11.7	2.8	17.8
Chad	12.2	6.0	11.5
Nita	10.5	2.2	19.4
Amie	10.3	2.9	17.5
Rae	12.7	2.4	16.9
Darcy	11.0	2.9	20.8
Lakisha	5.7	2.9	14.8
Lilia	6.3	3.1	23.0
Ralph	13.1	3.0	19.0
Andrew	5.8	2.9	6.5
Jacob	5.5	2.3	17.6
Mahwane	5.9	3.2	12.0
Robert	9.5	3.7	10.1
Sue	6.9	4.1	11.4
Kristen	4.1	2.0	18.5
Vivian	10.0	3.0	17.0
Dennis	8.6	2.6	13.3
Betty	6.4	2.2	6.5
Allen	8.3	2.0	15.1
Alex	7.2	6.4	12.3
Clara	10.4	4.7	10.3

Which walking rate was best?

Closeness Measure (show work):

Consistency Measure (show work):

Unit Quiz

Exploring Generalization Unit 4

1. Celo and Lexus each measured the length of the school's parking lot 30 times. Celo used a walking wheel, which is not a very precise tool. Lexus used a very precise laser to measure the length each time. Circle below what you expect to be true about their data.

   (a) The means of their measurements will be:

   very different **about the same** *(circle one)*

   Because:

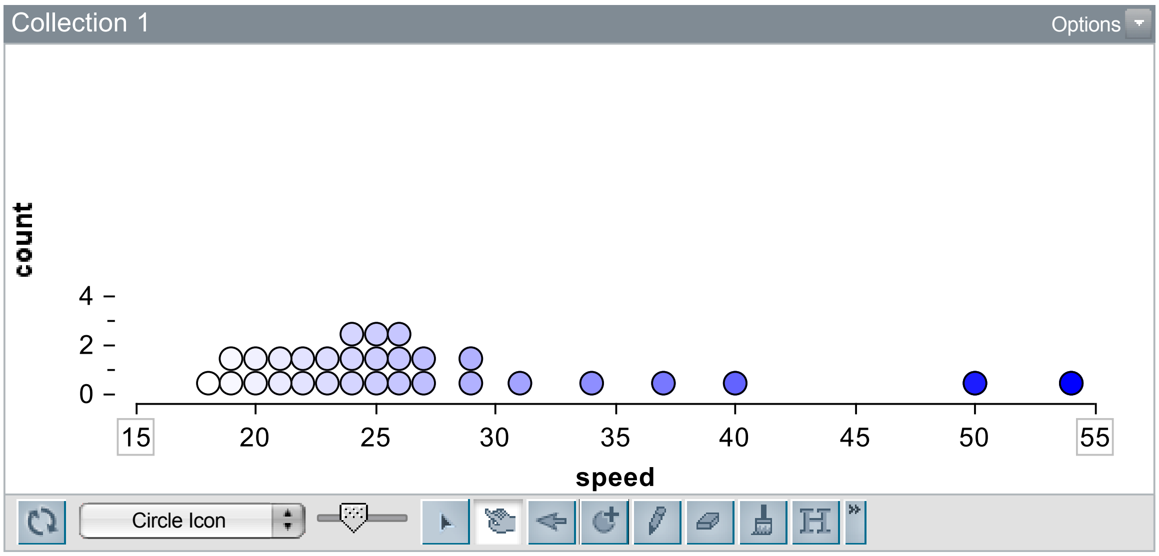
   (b) The average deviations of their measurements will be:

   very different **about the same** *(circle one)*

   Because:

2. The speeds of 30 cars were tracked by radar in a traffic zone.
Open SpeedLimit.tp to see the speeds, or use the data below.

25, 25, 19, 27, 22, 24, 23, 20, 37, 26, 23, 19, 25, 26, 22, 40, 54, 24, 26,
18, 50, 31, 27, 29, 21, 34, 29, 21,24, 20

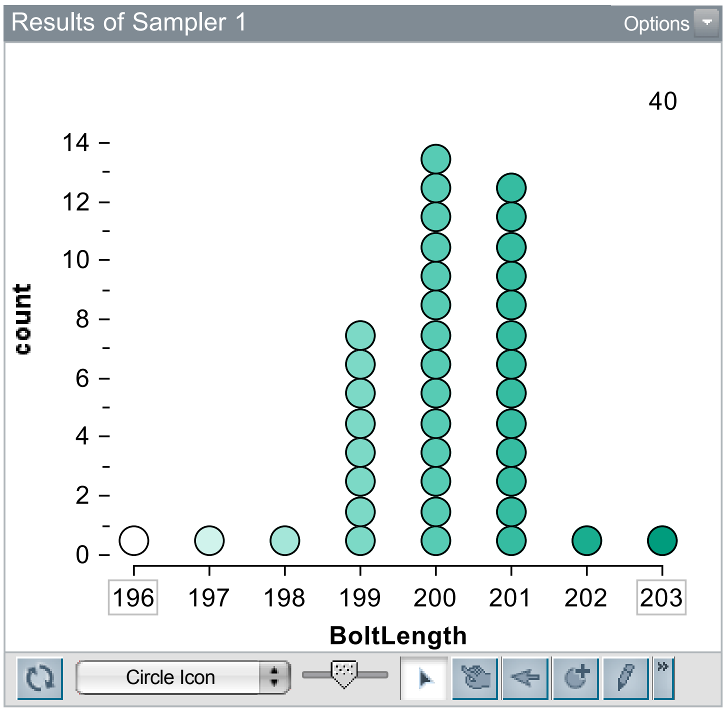


Choose a statistic to estimate the speed limit. Why did you choose that statistic?

Choose a statistic to estimate how consistently drivers tended to follow the speed limit. Why did you choose that statistic?

3. Acme Machines makes 200 mm long bolts for the engines of cars. Car makers require that the bolts be no longer than 202 mm or shorter than 198 mm. If more than 5% of Acme's bolts fall outside of this range, then Acme loses its contract.

   Here's a sample of 40 bolts from the Acme production line:



Should Acme keep its contract?

YES **NO** *(Circle one)*

Support your decision with mathematical reasoning.

Scoring Guide, CoS: Question 1(a), Celo and Lexus and Means of Measurements		
Levels	**Response Exemplars**	**Example of Student Response**
CoS3E	Predict the effect on a statistic of a change in the process generating the sample.	• "Even though they are using different tools, they are still measuring the same thing so both of their answers should be around the actual measurement." • The tools are different but the parking lot is not changing, so the mean is the same." • "They should be close together if they measured the same thing." • "They might be a little difference but not very much, they should be close."
CoS3C	Generalize statistic. (Might misunderstand but is thinking about generalizing.)	• "If Celo used those walking wheels, then there would be 30 numbers added up but none of them would be the length so would mess up the answer a lot, while Lexus has a more precise measuring tool that would probably not mess up his answer."
CoS 2 A	Calculates statistic of center.	• "To find the mean, you add them up and divide by the number of measurements."
NL(ii)	Response is relevant but unclear. Student may also answer yes or no without providing an explanation.	• "They did it 30 times." • "They are different."
NL(i)	Response is irrelevant, unclear, or a restatement of given information.	• "I'm not sure."
M	Missing response	

Scoring Guide, CoS: Question 1(b), Celo and Lexus and Average Deviations		
Levels	**Response Exemplars**	**Example of Student Response**
CoS3E	Predict the effect on a statistic of a change in the process generating the sample.	• " The average deviation is measuring precision, and Celo's tool is not very precise. How measurements will probably differ more, because his tool, again, is not very precise." • "Different because the measurements are more clumped up with the laser tool." • "With a precise tool the extremes won't be that extreme but the non-precise tool will have big extremes which will throw off the AD."
CoS3C	Generalize statistic.	• "If they measured the same or correctly, then they should have the same average deviation."
CoS2B	Calculate a statistic of variability.	• Student explains how to calculate an average deviation. • "If their median turns out to be different it messes everything up and the average deviation won't be close together."
NL(ii)	Response is relevant but unclear. Student may also answer yes or no without providing an explanation.	• Student says they will be about the same or different but does not offer any rationale.
NL(i)	Response is irrelevant, unclear, or a restatement of given information.	• "I'm not sure."

M	Missing response	

Scoring Guide, CoS: Question 2(a), Speed Limit		
Levels	**Response Exemplars**	**Example of Student Response**
CoS3 F	Choose/Evaluate statistic by considering characteristics of the sample.	• "The median. Not the mean because of the 2 very fast speeds."
CoS3 C	Generalize the use of a statistic beyond its original context of application.	• "I choose the median because if you wanted to find the speed limit, it will be a middle number. And 25 is a speed limit I have seen." • "I chose the mean because it is the average speed that the cars drove at. And it makes sense for a speed limit to be about 27."
CoS2 A	Calculate statistics indicating center	• Student calculates or estimates a mean or median, but does not explain why. • "25 is a mode. So that's it."
CoS1	Describe qualities of distribution informally.	• "Most of the speeds look around 25 mph, so that it what I chose."
NL(ii)	Response is relevant but unclear.	• "Speed limits are usually around 30 mph."

NL(i)	Response is irrelevant, unclear, or a restatement of given information.	• "I'm not sure."

Scoring Guide, CoS: Question 2(b), Speed Limit		
Levels	**Response Exemplars**	**Example of Student Response**
CoS4A	Predict that a statistic's value will change from sample to sample.	• "You don't want the range because it could change a lot if they had another sample."
CoS3F	Choose/Evaluate statistic by considering characteristics of the sample.	• "I chose the average deviation, because it shows precision and how much the speeds deviate from the average. The range is just two values."
CoS3C	Generalize the use of a statistic beyond its original context of application.	• Student chooses average deviation, IQR or range as measures of consistency of speed.
CoS3A	Invent an idiosyncratic measurement.	• "17/40 are near 25." • "20 to 25, there are more dots there."
CoS2 A	Calculate statistics indicating center. (The student confuses measures of center with measures of consistency/variability.)	• "I would choose the mean because it has fair shares." • "I chose the mode because it will show how precisely the drivers tended to follow the speed limit. The range is 26."
CoS1	Describe qualities of distribution informally.	• "Most of the speeds clump around 25 mph."
NL(ii)	Response is relevant but unclear.	• "People who speed do not drive consistently."
NL(i)	Response is irrelevant, unclear, or a restatement of given information. Student doesn't compare the real data to	

Unit Quiz Scoring Guides

Scoring Guide, CoS: Question 3, Acme Machines		
Levels	**Response Exemplars**	**Example of Student Response**
CoS4 A	Predict that a statistic's value will change from sample to sample.	• "It's more than 5% but it could be just this sample. Next time it could be just 1 or 2" • "It's close and it might be different next time."
CoS3C	Generalize use of a statistic.	• "If every dot is 2.5% then 7.5% of the bolt production is out of the range of sizes. Maching Acme lose their contract." • "Yes, it's only 3/40. That's close." • "No, they shouldn't keep the contract because 5% is 2 numbers out of range but they have 3" • The IQR is 2 (but might miscalculate) so they are mostly around 200."
CoS2B	Calculate statistics indicating variability.	• "The IQR is 2." (no explanation or interpretation)
CoS2A	Calculate statistics indicating center	• "The median is 200." (No explanation or interprets mdn as consistency.)
CoS1	Describe qualities of distribution informally.	• "Because He has the bolts right in the middle. Mostly all of his bolts are no longer than 202 mm and not shorter than 198 mm He has them right in the middle (DaD 3 response)."
NL(ii)	Response is relevant but unclear.	• Student suggests that consistency is not the same as the target value/
NL(i)	Response is irrelevant, unclear, or a restatement of given information. Student doesn't compare the real data to	