Learning to work with bivariate data, a key goal of middle-grades statistics curricula, is aided by a sequence of lessons.

WORKING WITH

n i o

IN

se

BIVARIATE DATA

Randall E. Groth,
Matthew Jones, and
Mary Knaub

Consider the three scatterplots shown in **figure 1**. Each graph shows students' test scores and the amount of time they spent studying. Looking at the graphs, approximately how long do you think one would have to study to get an 80 on each test? How long to obtain a score of 100? Now, put yourself in the position of a student who has not yet studied best-fit lines. How would you make predictions for each data set? What unique difficulties might you have with each one?

The graphs shown in **figure 1** differ in the amounts and types of noise they contain. We use the word *noise* to refer to statistical variability. Statistics has been characterized as "the study of noisy processes—processes that have a signature, or signal, we can detect if we look at sufficient output" (Konold and Pollatsek 2002,

p. 260). The graph shown in **figure 1a** has a more pronounced linear "signal" than the others. The **figure 1b** scatterplot has a similar amount of noise, but it also contains a different type of noise—a data point that diverges a great deal from the otherwise linear signal portrayed by the other points. **Figure 1c** is the noisiest of the three, but one might still be able to discern a linear signal after a close examination.

We asked a group of students during a classroom research study to analyze data sets containing different amounts of noise. The four students in our group were finishing seventh grade and participating in summer mathematics instruction. We refer to them using the pseudonyms Rebecca, Shonice, Joseph, and DuJuan. We sought to design and test ways to prepare them for statistics in eighth grade, when they would need to discern the following:

> Know that straight lines are widely used to model relationships between two quantitative variables. For scatter plots that suggest a linear association, informally fit a straight line, and informally assess the model fit by judging the closeness of the data points to the line. (CCSSI 2010, p. 56)

Research has shown that these ideas can be quite challenging (Casey 2015; Pérez and Daiga 2016). Therefore, we carefully examined how our small group of students made sense of data sets with varying amounts and types of noise to discover the students' reasoning patterns so that we could design suitable instruction.

Fig. 1 These data sets show the number of minutes spent studying and corresponding test scores for three different tests.

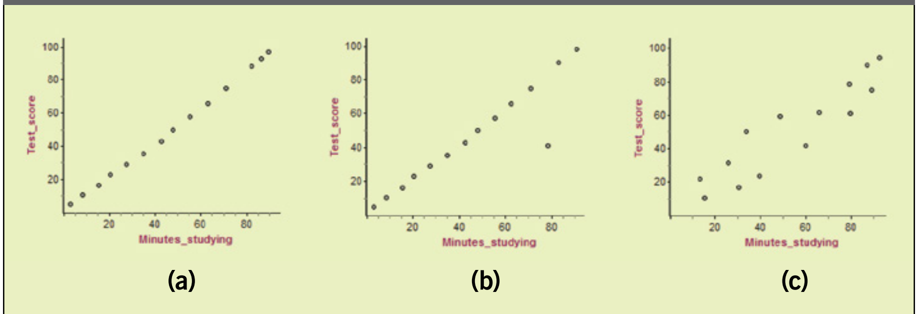


Fig. 2 Rising eighth graders created this Spaghetti Bridges poster after their first summer statistics lesson.

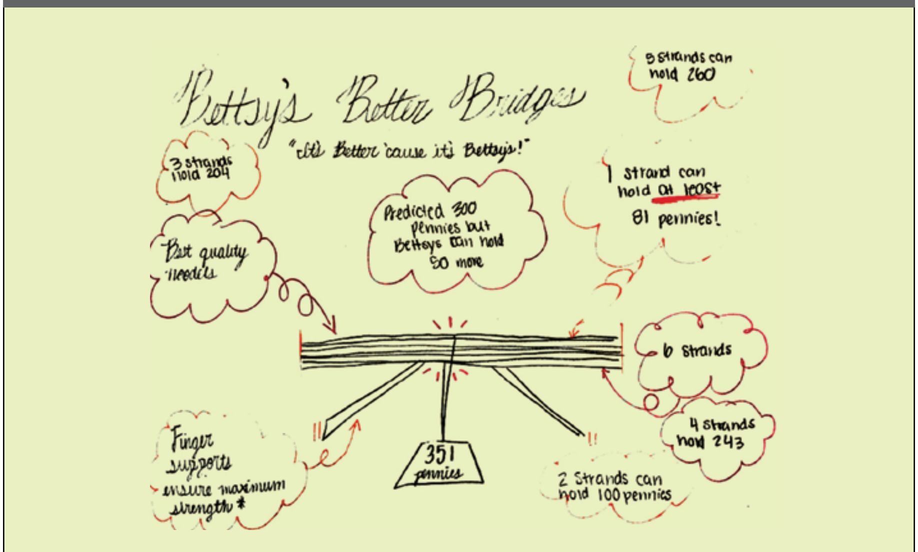
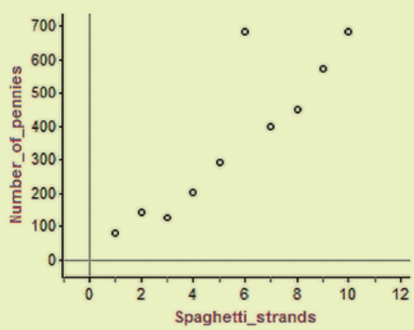


Fig. 3 This data set shows the numbers of pennies held by different numbers of spaghetti strands.

Spaghetti_strands	Number_of_pennies
2	145
3	126
4	205
5	293
6	684
7	397
8	451
9	575
10	682



MODERATE NOISE: SPAGHETTI BRIDGES

We taught seven one-hour weekly sessions during the summer. During the first two sessions, students used spaghetti strands to represent bridge beams and used pennies in a cup to represent the weight on a bridge (Kroon 2016). The cup had holes cut on either side near the top to allow for the insertion of several spaghetti strands. Students found how many pennies it would take to break one spaghetti strand and then conducted several more trials, adding a strand each time. The number of strands was one variable, and the number of pennies was the second variable. Rebecca and Shonice created a poster

Fig. 4 Rebecca used an oval to highlight the aggregate in a scatterplot.

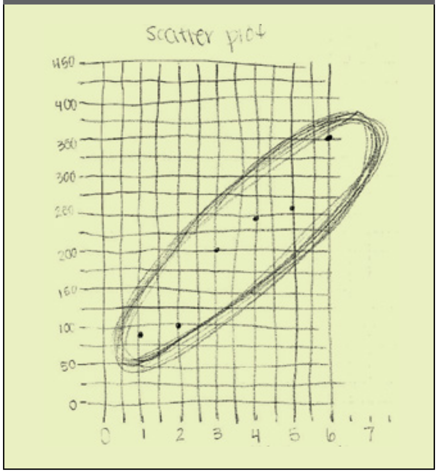


Fig. 5 Shonice's description of the Fruit-by-the-Foot graph led her researchers to believe that she viewed the data as an aggregate.

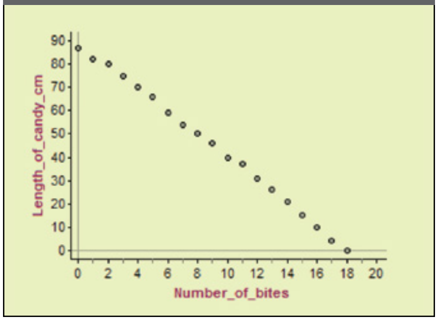
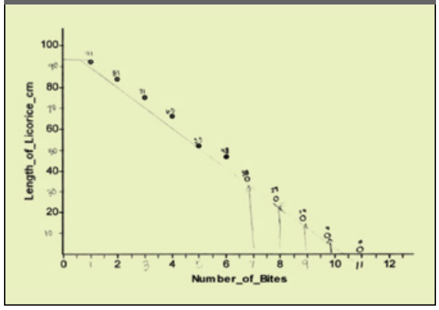


Fig. 6 Rebecca attempted to fit a line to bivariate data.



to represent their work during the first lesson (see **fig. 2**).

One of the spaghetti-bridge data sets used for class discussion is shown in **figure 3**. After seeing the scatterplot, Shonice and Rebecca said that it reminded them of a connect-the-dots pattern. Shonice went up to the board where the graph was displayed and motioned how she would connect the dots to one another; students' initial inclination was not to fit a straight line to the data to discern the trend. We did not correct students at this point to tell them they should fit a straight line rather than connect the dots. Our goal was for the straight-line representation for "signal" to emerge naturally, as students made sense of data sets, rather than having the tool imposed on them by the teachers.

To see straight-line patterns in bivariate data, students must perceive the data as being one entity, or an aggregate, rather than just a series of individual points (Bakker and Gravemeijer 2004). The "noisiest" point in the **figure 3** scatterplot, at (6, 684), helped us draw students' attention to the aggregate. We asked students what a good name for the point would be. Some of the names they gave it were "outcast," "outsider," and "out-of-order." In response, we introduced the formal term *outlier* and drew an oval around the main cloud of points, leaving (6, 684) outside the oval. We asked students to describe the trend of the points inside the oval. Rebecca motioned upward in a straight line from left to right on the graph and said the pattern was "climbing to the top." Joseph observed, "The more spaghetti strands, the higher the number of pennies." Hence, the noise from the outlier was valuable in helping draw attention to the aggregate and its upward linear trend. Rebecca also later used this oval representation to highlight the aggregate when working with a different data set (see **fig. 4**).

Some of the names they gave the point were "outcast," "outsider," and "out of order," at which time we introduced the formal term *outlier*.

MINIMAL NOISE: FRUIT-BY-THE-FOOT™

Although students had begun to focus on the aggregate while working with spaghetti bridges, sometimes they still examined individual data points when it would have been better to consider the larger picture. Shonice, for example, disagreed with Joseph's observation that the number of spaghetti strands and the number of pennies increased together, since three strands of spaghetti held less weight than two strands (see **fig. 3**). To better focus students' attention on the overall trend, during our next two lessons, we had them gather data from a context with a minimal amount of noise. Each student received a strand of Fruit-by-the-Foot. They took bites from the strand. After each bite, they recorded the remaining length of the strand. They plotted their data in a scatterplot, with number of bites on the horizontal axis and remaining length on the vertical axis. We hoped

Fig. 7 Test score data: The *x*-variable represents the number of hours studied, and the *y*-variable represents the score on a final test.

x	10	12	14	14.5	15	16	16	17	17	18	18	18	19	19	20	22	22	22	23	25	26	27	28	29.5	30
y	47	53	59	63	61	57	59	76	72	59	63	71	60	61.5	58	74	76	84	81	75	34	?	89	96	92

Fig. 8 Students were not naturally inclined to use straight lines to make their predictions; Rebecca (graph on left) and Shonice (graph on right) used connect-the-dots representations.

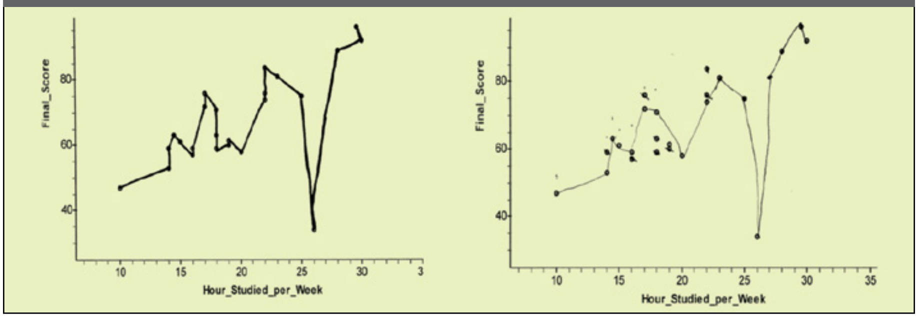
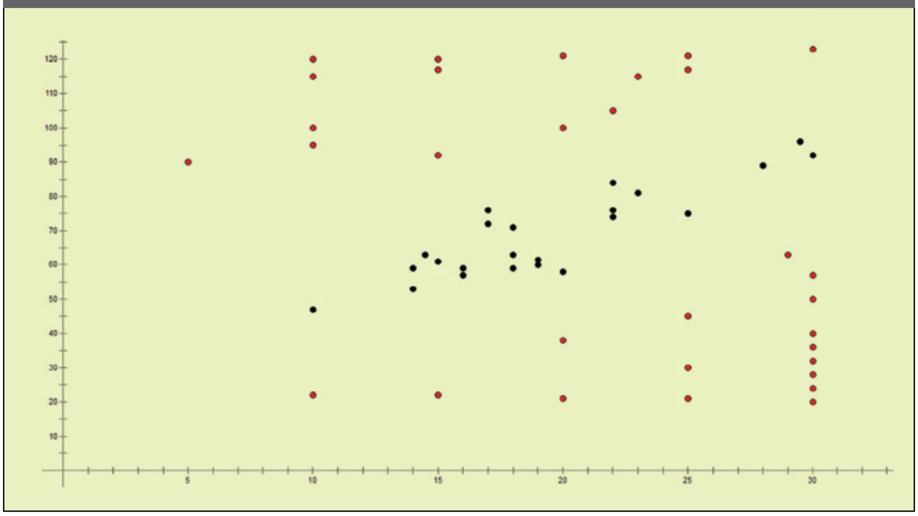


Fig. 9 Students indicated unreasonable predictions by marking (in red) various numbers of hours spent studying.



that the resultant scatterplots would help students focus on the overall linear trend of the data and understand negative linear relationships.

During the Fruit-by-the-Foot activity, as expected, students produced graphs with pronounced linear signals and very little noise. **Figure 5**, for example, shows Shonice's data. As she produced the graph, Shonice remarked how straight it was, hence attending to an aggregate feature. When we asked her about the relationship in the graph, she said that as the number of bites went up, the length of the strand went down. She motioned downhill left to right to indicate a negative relationship. Shonice's comments and actions suggested that she was viewing the data as an aggregate rather than just a collection of individual points.

Because our students commented on the straightness of the graphs they produced, we asked them to sketch straight lines through their data to capture the trends they saw. For example, Rebecca attempted to fit a straight line to one particular data set (while thinking about the similar situation of taking bites of licorice) (see **fig. 6**). When we asked students what the lines meant and how they might be useful, Shonice said they could help find how many bites it would take someone to finish eating a strand. When asked how the graphs compared to the spaghetti bridges graphs, she motioned uphill from left to right to illustrate the trend in the spaghetti-bridge graph and downhill from left to right for the Fruit-by-the-Foot. The other students used similar reasoning in contrasting the two types of data sets. Seeing a graph with a clear linear signal and contrasting positive and negative linear relationships appeared to have helped students employ aggregate reasoning and discern trends. The Fruit-by-the-Foot experiment had also provided a natural context for fitting straight lines to data.

CONSIDERABLE NOISE: TEST SCORES

Given the apparent success of the Fruit-by-the-Foot activity, we decided to see how students would handle a data set with considerably more noise. We introduced data (see **fig. 7**) that showed the scores obtained on a major test (we compared it with a medical license exam) after different amounts of time spent studying. We asked students to describe the overall trend shown in the data and to predict the score that one could expect after studying for 27 hours.

The noisier data set caused more trouble than we had anticipated. Students did not immediately represent the data in a scatterplot to discern the trend. When given a scatterplot, they

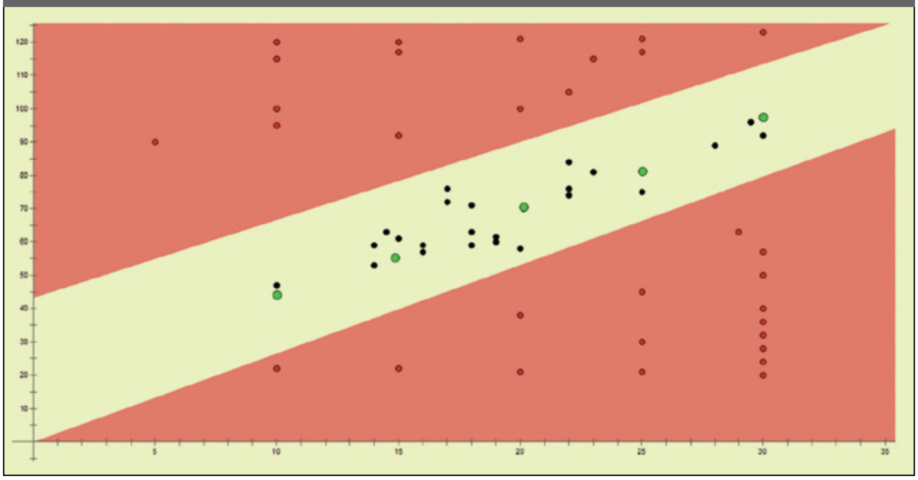
As we continued to ask students to make predictions for missing data values, they were not naturally inclined to use straight lines to make their predictions.

-M-I-S-H-A-/THINKSTOCK

did not initially fit straight lines to the data to help justify their predictions. Instead, Rebecca and Shonice reverted to connecting all the dots in trying to describe its trend (see **fig. 8**). Rebecca began to doubt the positive relationship between the two variables because they did not increase uniformly. The outlier (26, 34) unduly influenced Joseph's estimate of the score that one could expect after studying for 27 hours. DuJuan employed point-by-point reasoning in making his estimate rather than looking at overall trend. In response to these difficulties, we put an oval around the main cloud of points (as in **fig. 4**) and suggested fitting a line to the data. However, as we continued to ask students to make predictions for missing data values, they were not naturally inclined to use straight lines to make their predictions.

The approaches we had used for previous lessons appeared to be insufficient for the noisier test score data. We wanted students to perceive the straight lines fitting to data as valuable tools for making reasonable predictions. However, just pointing out that lines can be used for this purpose had not been very effective. So, in our final lesson, we decided to have students contrast *reasonable* and *unreasonable* statistical predictions. We believed that contrasting these concepts with one another could be helpful, just as contrasting an outlier with the main cloud of data had helped bring both of those concepts into sharper relief.

Fig. 10 Students' most reasonable score predictions are marked in green for different numbers of hours studying.



To begin our final lesson, we asked students to examine the test score data again, but this time with the outlying value of (26, 34) removed. We told students that the outlier was actually a mistake made in recording the data. Hence, we were justified in deleting it because it represented a data recording error rather than an actual data point. Pedagogically, we actually made the decision to remove the outlier to at least temporarily reduce the amount of noise that students had to deal with in the data. With the modified data set projected on a whiteboard, we asked students to come up and take turns marking in red *unreasonable* test score predictions for various amounts of hours spent studying. They produced unreasonable predictions (see **fig. 9**). We asked students to shade in red the entire region containing the unreasonable points.

Next, we asked students to come to the board and make several reasonable predictions for various numbers of hours spent studying. They marked these points in green on the scatterplot. All the students stayed within the middle region of the graph, not shaded in red, when making their predictions. Students predicted values near the center of the unshaded region, with the exception of DuJuan, who put his predictions closer to the top of it. With several reasonable predictions on the board, we asked students to narrow down the *most* reasonable predictions for different numbers of hours spent studying. Students took turns marking the most reasonable score predictions for 10, 15, 20, 25, and 30 hours spent studying. Students' predictions tended to fall near the center of the unshaded region (see the green points in **fig. 10**). With the green points shown in

figure 10 on the classroom whiteboard, we asked students to describe any patterns they observed in the data. Rebecca came up to the board and connected the green points. Joseph noticed that she had produced a nearly straight line. We asked him to use a meterstick to sketch a straight green line to show this linear pattern. His line went through the middle of the unshaded region and was near each of the green points on the board (see **fig. 11**). DuJuan called Joseph's line a "base" to show "where the points were going." The process of isolating a reasonable region for predictions and focusing on the center of the reasonable region helped our students identify a linear signal in the noisy test score data set.

To conclude our final lesson, we wanted to see if students could use the linear signal they had identified to make predictions for missing data values. We chose several numbers of hours spent studying that had not been in the data set and asked students to predict test scores for each one. For example, we had the following exchange with DuJuan:

Teacher: I took this test, and I studied for 32 hours. Can you show me what score you'd think I would get with an X on the graph?
DuJuan: 32 hours?
Teacher: 32 hours.
DuJuan: [Takes a few moments to study the scatterplot.] Around 80.
Teacher: Could you mark it with an X on the graph up there?

DuJuan went to the whiteboard and plots a point on Joseph's best-fit line (shown in **fig. 11**) with 32 as its x-coordinate.

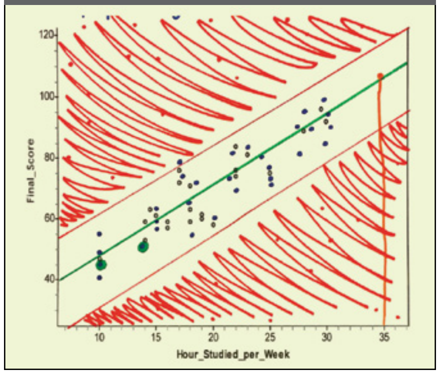
Teacher: Alright, how did you figure that one out?
DuJuan: It's the trend.
Teacher: The trend? So, what showed you the trend?
DuJuan: The green line.

It is interesting to note that DuJuan was the only student who did not consistently place his predictions directly on the green informal best-fit line we had produced, even though he did so in this case. Some of DuJuan's predictions fell on the line, but he went slightly above the line in cases where he believed one should get a higher score from several hours spent studying. All the predictions, however, fell within the reasonable region that emerged during class discussion, which placed them near the linear signal that we had identified.

TEACHING AMID NOISE

The amount and type of variability, or noise, in bivariate data was not identified as a consideration in the standards we sought to help students attain (CCSSI 2010). However, we found that it had a strong impact on students' thinking. Because of this, noise became an important consideration as we designed lessons for students. We found that we could sometimes use noise to our advantage, such as when we used outlying points to draw attention to aggregate patterns in spaghetti-bridge data (see **fig. 3**). At other times, we had to help students cut through the noise to identify a linear signal in bivariate data, as with the test score data (see **fig. 7**).

Fig. 11 Rebecca depicted Joseph's trend line.

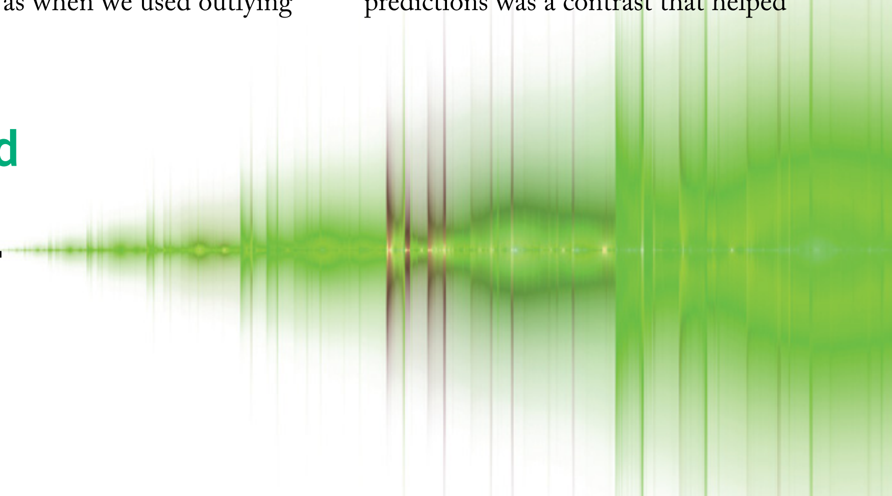


We used noise to our advantage in helping students perceive bivariate data as an aggregate. Specifically, encountering outliers helped spark discussion about how they differed from the other data points. During these discussions, we talked about how the outlier fell outside the main cloud of data. It then became natural to talk about the main cloud of data as a single entity, or an aggregate. Students began identifying this main cloud of data by enclosing it with an oval (see **fig. 4**). So, drawing a contrast between the outliers and the rest of the data set helped students begin to discern trends and patterns within the aggregate.

Reasonable versus unreasonable predictions was a contrast that helped

IGORKIRILLOV/THINKSTOCK

Encountering outliers helped spark discussion about how they differed from the other data points.



our students cut through noise in bivariate data to detect a linear signal. As students explored this contrast in the context of test score data, a reasonable region emerged (see **fig. 10**). The center of the reasonable region became a natural place for students to put an informal line of best fit to use for making predictions about data values not included in the set (see **fig. 11**). Drawing this contrast for students was important for the noisy test score data set, since they did not identify its linear signal as readily as they did in less noisy data sets, such as those generated in the Fruit-by-the-Foot context (see **figs. 5** and **6**).

We hope that our classroom experiences help other teachers work with noise as a factor in students' learning of bivariate data analysis. We have outlined one possible path and set of tasks to use for facilitating classroom discussions. Readers may wish to try our strategies with their own students and devise some of their own. As our students become more adept at finding signals in noisy processes, they simultaneously become more familiar with one of the most fundamental aspects of statistical thinking.

ACKNOWLEDGMENT

The work described in this article was supported by the National Science Foundation (NSF) under Award No. DRL-1356001. The views expressed are those of the authors and do not necessarily reflect those of the NSF.

REFERENCES

Bakker, Arthur, and Koeno P. E. Gravemeijer. 2004. "Learning to Reason about Distribution." In *The Challenge of Developing Statistical Literacy, Reasoning, and Thinking*, edited by Dani Ben-Zvi and Joan Garfield, pp. 147–68. Dordrecht, The Netherlands: Kluwer Academic Press.

Casey, Stephanie A. 2015. "Examining Student Conceptions of Covariation: A Focus on the Line of Best Fit." *Journal of Statistics Education* 23 (1). www.amstat.org/publications/jse/v23n1/casey.pdf

Common Core State Standards Initiative (CCSSI). 2010. Common Core State Standards for Mathematics. Washington, DC: National Governors Association Center for Best Practices and the Council of Chief State School Officers. http://www.corestandards.org/wp-content/uploads/Math_Standards.pdf

Konold, Clifford, and Alexander Pollatsek. 2002. "Data Analysis as the Search for Signals in Noisy Processes." *Journal for Research in Mathematics Education* 33 (July): 259–89.

Kroon, Cindy D. 2016. "Spaghetti Bridges: Modeling Linear Relationships." *Mathematics Teaching in the Middle School* 21 (May): 563–68.

Pérez, Arnulfo, and Michael Daiga. 2016. "Data Analysis, Statistics, and Probability." In *What Mathematics Do Students Know and How Is That Knowledge Changing? Evidence from the National Assessment of Educational Progress*, edited by Peter Kloosterman, Doris Mohr, and Crystal Walcott, pp. 179–210. Charlotte, NC: Information Age Publishing.



Randall E. Groth, regroth@salisbury.edu, teaches mathematics education courses at Salisbury University in Salisbury, Maryland. He is interested in designing ways to help develop students' statistical thinking. **Matthew Jones**, mjones9@gulls.salisbury.edu, and **Mary Knaub**, mknaub0022@live.worwic.edu, were selected to be participants in a National Science Foundation (NSF) Research Experiences for Undergraduates (REU) project at Salisbury University. They collaborated on the design and implementation of the study reported in this article.

Let's Chat about Data Noise

On Wednesday, October 18, 2017, at 9:00 p.m. ET, we will expand on "Working with Noise in Bivariate Data" (pp. 82–89), by Randall E. Groth, Matthew Jones, and Mary Knaub. Join us at #MTMSchat.

We will also Storify the conversation for those who cannot join us live. The *MTMS* monthly chats fall on the third Wednesday of the month.