MATHEMATICS TEACHER
Learning & Teaching
PK-12

Mathematics Teacher: Learning and Teaching PK-12, is NCTM's newest journal that reflects the current practices of mathematics education, as well as maintains a knowledge base of practice and policy in looking at the future of the field. Content is aimed at preschool to 12th grade with peer-reviewed and invited articles. *MTLT* is published monthly.

ARTICLE TITLE:
Considering the 5 Practices Through a Statistical Lens

AUTHOR NAMES:
Smucker, Karoline; Sepúlveda, Francisco; Weiland, Travis; Cannon, Susan; Casey, Stephanie; and Byun, Sunghwan

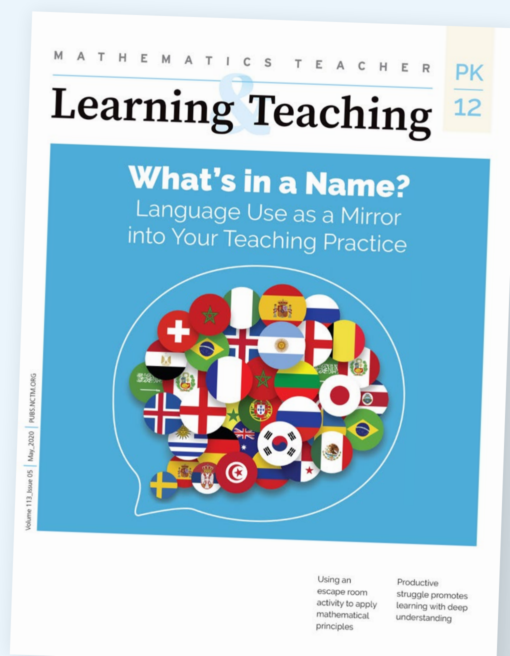
DIGITAL OBJECT IDENTIFIER: 10.5951/MTLT.2023.0233
VOLUME: 117
ISSUE NUMBER: 2

Mission Statement

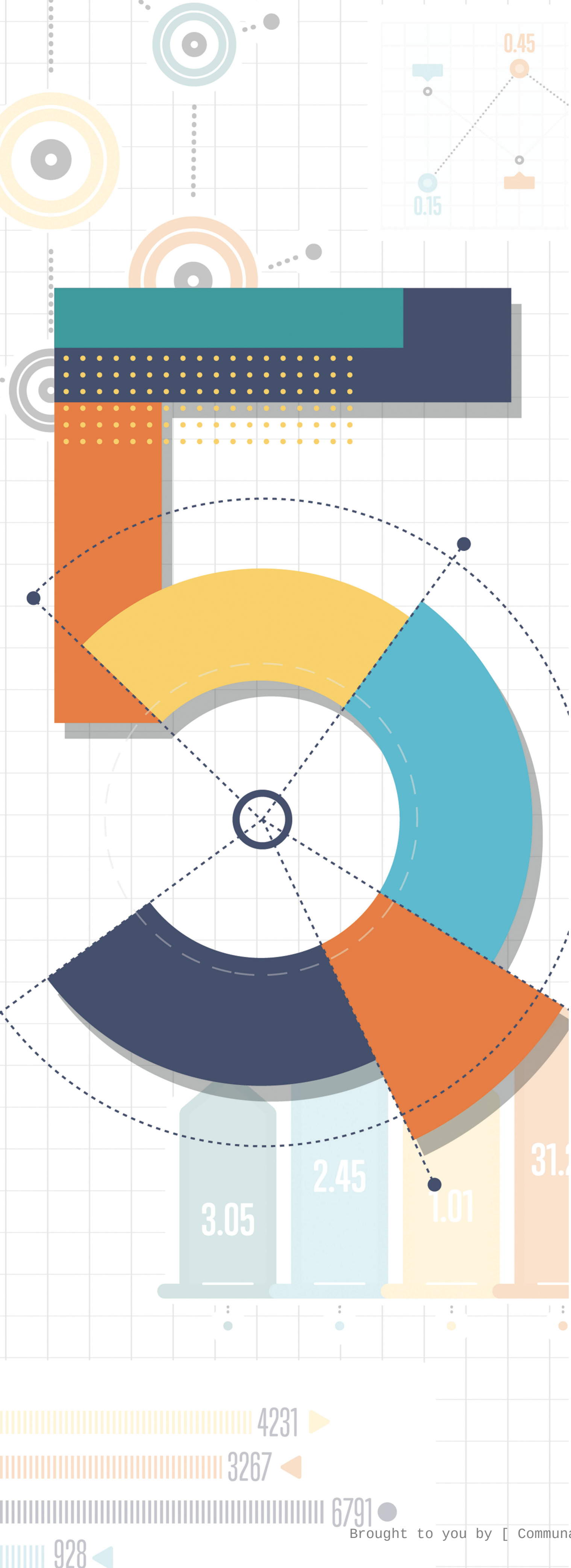
The National Council of Teachers of Mathematics advocates for high-quality mathematics teaching and learning for each and every student.

Approved by the NCTM Board of Directors on July 15, 2017.

CONTACT: mtlt@nctm.org



NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS



Access digital content at nctm.org/mtlt11702fc.

Considering the 5 Practices Through a Statistical Lens

An adaptation of The 5 Practices framework for statistical investigations that accounts for differences between mathematics and statistics.

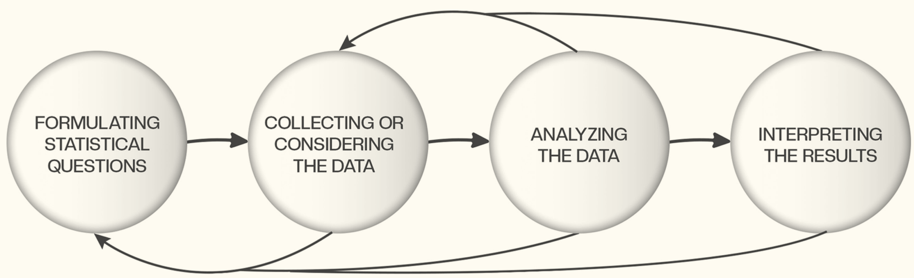
Karoline Smucker, Francisco Sepúlveda, Travis Weiland, Susan Cannon, Stephanie Casey, and Sunghwan Byun

Statistics has been a content domain in the mathematics curriculum for decades (National Council of Teachers of Mathematics [NCTM], 2000). However, statistics is a distinct discipline from mathematics and there are important differences in how one should teach the two disciplines (Cobb & Moore, 1997). In this article, we consider how these differences can inform an *adaptation* to the 5 Practices for Orchestrating Productive Mathematics Discussions framework (Smith & Stein, 2018) to engage students and teachers in meaningful statistical investigations and the ensuing class discussions. The 5 Practices are a key pedagogical tool for building productive discourse in mathematics classrooms featured in NCTM's (2014) *Principles to Actions* document. The practices for orchestrating productive mathematics discussions are planning, anticipating, monitoring, selecting, sequencing, and connecting (Smith & Stein, 2018). Others have described their use of the practices in teaching statistics with preservice teachers (Groth, 2015).

We begin by noting three key differences between mathematics and statistics that need to be recognized in the teaching of the two disciplines: (1) consideration of context, (2) the omnipresence of variability, and (3) inference. When doing mathematics, students can operate in complete abstraction; this is not possible in statistical inquiry. In statistical inquiry, according to the Guidelines for Assessment and Instruction in Statistics Education II (GAISE II), students should participate in an investigative process (see Figure 1) that includes (a) formulating statistical questions, (b) collecting or considering the data, (c) analyzing the data, and (d) interpreting the results (Bargagliotti et al., 2020).

The context of the investigation is crucial throughout the statistical investigation process and is key to statistical reasoning (Bargagliotti et al., 2020; Cobb & Moore, 1997; Wild & Pfannkuch, 1999). Context drives

Figure 1 GAISE II Statistical Problem-Solving Process



Note. Adapted from Bargagliotti et al., 2020.

Karoline (Karie) Smucker, she/her, ksmucker@eou.edu, is currently an assistant professor of education at Eastern Oregon University, after teaching high school mathematics for 12 years. She is particularly interested in how technologies like CODAP can impact the teaching and learning of statistics and probability.

Francisco Sepúlveda, he/him, francisco.sepulveda@cinvestav.mx, is currently pursuing a PhD at Cinvestav in Mexico City, focusing on the exploration of informal inferential reasoning for his dissertation. He also teaches statistics at an undergraduate level and enjoys tutoring students across all grades.

Travis Weiland, he/him, tweiland@uh.edu, is currently an assistant professor at the University of Houston and studies the teaching and learning of statistics and data science in K–12 settings through data investigations of sociopolitical issues.

Susan Cannon, she/her, cannonso@uga.edu, is an assistant professor of mathematics education at the University of Georgia. She taught mathematics and science in Atlanta, GA for 13 years prior to transitioning to higher education.

Stephanie Casey, she/her, scasey1@emich.edu, is a professor of mathematics education at Eastern Michigan University in Ypsilanti, MI. She dedicates much of her work to improving secondary teachers' preparation to teach statistics and data science.

Sunghwan Byun, he/him, sbyun2@ncsu.edu, is an assistant professor of mathematics education at North Carolina State University and researches discourse and social interaction for teaching and learning mathematics, statistics, and data science.

doi:10.5951/MTLT.2023.0233

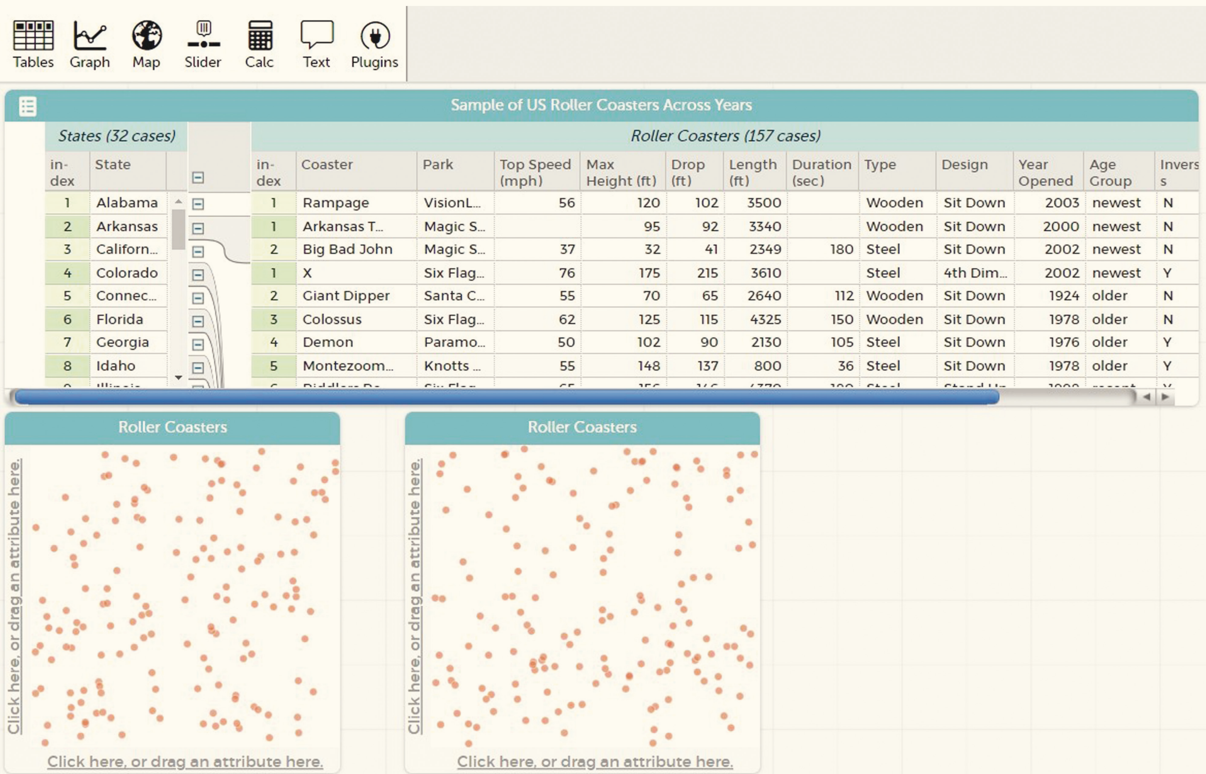
what data is collected, how it is collected, what questions are asked of it, how it is analyzed, and how the results can be interpreted. Therefore, teachers should consistently encourage students to connect to the data's context when doing statistical investigations. Selection of contexts that are meaningful to students and their community and/or connect to other disciplines can also be motivational for students (NCTM, 2014).

Another major difference in the disciplines and the primary reason for the development of statistical tools is the omnipresence of variability in data (Cobb & Moore, 1997). Exploring variability in data is crucial to statistical practice. When asking a statistical question, such as "How many teeth have students in third grade lost?," students should expect variability in the responses (i.e., they should not expect all students to have lost the same number of teeth). To answer the statistical question, students need to consider how to account for and represent that variability in the data within the context of the investigation. Because of the omnipresence of variability in data, students need to learn to be comfortable with uncertainty and understand how to measure and make sense of uncertainty in statistical inquiry (Cobb & Moore, 1997).

The third key difference between the disciplines we want to highlight is inference, a fundamental concept in statistics. Whereas in the school mathematics curriculum the main focus is often on deductive reasoning, in statistics, students are asked to make inferences about the population based on a sample. There is often not one right answer that results from this process, which is a major difference between the two disciplines. While study of formal inference typically does not occur until taking a college-level statistics class, secondary students are typically recommended to learn to make informal inferences that generalize beyond the data (Bargagliotti et al., 2020), considering uncertainty with data as evidence (Makar & Rubin, 2009).

Given these important differences between statistics and mathematics, we offer an adaptation of the 5 Practices framework for classroom sessions centered around statistical investigations. This recommendation is rooted in literature regarding statistics and its teaching, along with our experiences teaching statistics

Figure 2 Roller Coaster Task Initial CODAP Environment



across all grade bands. Similar to Franklin et al.'s (2015) discussion of mathematical practices through a statistical lens, in this article, we aim to show how the 5 Practices for orchestrating mathematical discussions can be translated to teaching statistical investigations.

CLASSROOM VIDEOS OF STUDENTS EXPLORING CHARACTERISTICS OF ROLLER COASTERS

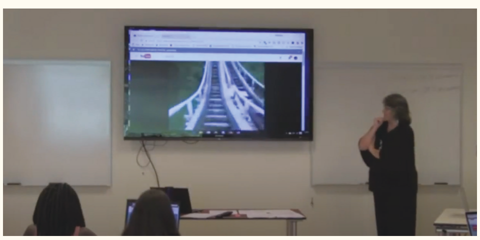
To give readers a view of what orchestrating a class session centered around a statistical investigation could look like, we use videos of an expert teacher leading a data-based investigation with a Grade 7 class in the U.S. These videos were created by the Enhancing Statistics Teacher Education with E-Modules (ESTEEM) project (link online). This exploration of roller coasters was framed using CODAP (link online), a free online data analysis platform (Concord Consortium, 2023). For more information about CODAP and how it may be used in teaching, see Mojica et al. (2019). The dataset used in the teaching episode consisted of 157 roller coasters with 15 associated variables, including categorical (e.g., the age group of the coasters) and quantitative variables (e.g., drop height—see Figure 2). Other variables included the maximum height, top speed, and length of each roller coaster. The data and the investigation platform can be accessed as an example document in CODAP (link online).

In the teaching episode, the teacher explains the stages of the statistical investigation to the students (pose, collect, analyze, and interpret). To engage students in the context, she then has the class watch and discuss a video that makes the students feel as if they are riding a roller coaster that the teacher rode when she was young (Video 1). After watching the video, the teacher leads a discussion of different characteristics of roller coasters that students care about, which leads into the launch of the task for the students to start investigating the maximum height of wooden roller coasters. Then, she asks the students to explore additional variables in the dataset to lead students to complete their own statistical investigations about roller coasters. Video 2 illustrates the interaction between the teacher and students for the remaining class time. The teacher circulates the classroom to monitor students' work and scaffolds students' investigation of the data using CODAP through feedback and suggestions. Toward the end of this video, the teacher leads a whole class discussion centered around student pairs presenting their work to the class. A detailed analysis of the videos can be found in the supplementary material (link online).

TRANSLATING THE 5 PRACTICES TO STATISTICS

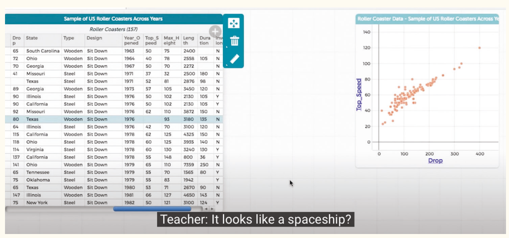
With this teaching episode in mind, we turn to the 5 Practices. The tables that follow provide an overview of the elements teachers should consider when adapting the 5 Practices framework to the context of statistical investigations. Essential questions for each practice are outlined in the left column. These questions are drawn from Smith et al. (2020). As a note, Smith and colleagues discuss solution strategies as the product of engaging in a cognitively demanding task, while in statistics, the end product of a statistical investigation is a well-reasoned argument using data to support claims. The right column showcases some key considerations when teaching a lesson centered around a statistical investigation task (in comparison with a typical mathematical task), as well as guiding questions for a teacher

Video 1 Students Watch and Discuss a Roller Coaster Video



Watch the full video online

Video 2 Students Perform a Statistical Investigation of Roller Coasters



Watch the full video online

to consider when planning or teaching the lesson. After each table, we describe how each practice can look. In addition, the supplementary material (link online) provides a more detailed overview of what each practice might look or sound like in a classroom setting. The first practice, Practice 0, involves planning for the task or investigation (see Table 1).

As teachers examine their learning goals and potential tasks for data investigations, a key feature of their planning will involve going beyond selecting tasks with high cognitive demand to consider datasets and digital platforms for exploring them. Decisions must be made about an appropriate data context, in particular, what data will be collected or considered, the size of the dataset, and the number of variables that are appropriate based on features like students' prior learning and whether technology will be used in the lesson. Teachers also must decide how to make the data context relevant to students, perhaps by investigating data that students collect themselves, or providing background on the context in some way to increase engagement. As an example from the selected teaching episode, when planning the lesson, in addition to obtaining the dataset about roller coasters, the teacher in Video 1 delved into contextual information such as the history and materials of roller coasters. Additionally, she found a video that helped her to introduce context and formulated an initial premeditated question to kickstart the inquiry: "What is the maximum height that wooden roller coasters tend to be?" The roller coaster context serves as an illustrative example. While this particular context will not be relevant to all learners, teachers can select data sets that connect to their students' backgrounds and interests. After planning the context and direction for the investigation, the teacher should turn to anticipating student approaches to the investigation, Practice 1 (see Table 2).

When using this practice in teaching mathematics, teachers must prepare for the variety of solution strategies that students will produce for the task. In teaching lessons centered around statistical investigations, teachers are required to anticipate potential graphical representations, analytical techniques, patterns observed based on the data context, and possible interpretations of the data. Anticipating which variables are likely to be investigated by students and the ways they may analyze these variables are a key component of this stage of planning. For example, in the selected teaching episode, the teacher anticipated attributes in the dataset that students were likely to be interested in about roller coasters (like type of material and maximum height) to purposefully bring those out in an

Table 1 Key Considerations and Questions for Planning a Data Investigation

Practice 0: Planning **Setting learning goals**: Does the goal focus on what students will learn about [statistics] as opposed to what they will do? **Selecting high-level tasks:** Does your task provide students with the opportunity to think, reason, and problem solve? What resources will you provide to ensure that all students can access the task? What will you take as evidence that students have met the goal through their work on the task? (Smith et al., 2020, p.14)	**Key considerations** • Selecting the context of the data to drive investigations (i.e., selecting issues that are relevant to your students and learning about that context more yourself). • Selecting data size (i.e., size of data set, number of variables). • Determining what part(s) of a statistical investigation will be the focus and which big statistical ideas align to learning goals. • Determining how data will be obtained (i.e., students collect, use pre-existing data set, teacher gathers and cleans). **Guiding questions for the teacher (during *planning*)** • What context would be relevant for students? How does it intersect with student interest? • What type of data do you need to meet this learning goal? • What do you know about the context of the data? How do you build student understanding of context and how do students draw on their experiences to help bolster their interpretation? • Who will collect data or will you consider data that has already been collected? • What population will the data come from? How much data should be obtained or collected? • What (if any) technology will you use? What are the affordances and constraints of that technology?

engaging way. Another important consideration for teachers when teaching in statistical investigations is learners' perspectives on data. Research by Konold and colleagues (2015) provides meaningful insight into ways learners view data that is useful for anticipating student approaches during statistical investigations. For example, data may be recognized as *case values*, with analysis focused on individual values and the attributes represented, while an *aggregate perspective* is indicated when students view data as a "unity with emergent properties such as center and shape" (p. 308). Konold and colleagues note that because statistics is "fundamentally about the behavior of aggregates," statistics educators must consider how to "foster students' abilities to perceive and reason" about data holistically (p. 302). In addition, when students investigate data that they have collected themselves, the teacher may need to be prepared to handle differences between the collected data and what they expected. In the supplementary material (link online), we highlight some specific ways this may look during data investigations. Once the teacher has anticipated representations and patterns that may emerge in the analysis phase of the investigation, they can turn to monitoring students during the enactment of the lesson (see Table 3).

The 5 Practices approach emphasizes documenting student strategies to the given task and the questions teachers ask of students to advance and assess their thinking (NCTM, 2014; Smith & Stein, 2018). Some specific questions teachers might use to advance and assess can be found in the supplementary material (link online). An important difference that teachers must consider when implementing data investigations is that not only may students end up investigating a variety of questions leading to different conclusions, but they may also pursue multiple graphs and analyses as part of their work in making sense of the data. Teachers may need to monitor not only the final product of the

Table 2 Key Considerations and Questions for Anticipating During a Data Investigation

Practice 1: Anticipating **Getting inside the problem:** How do you solve the task? How might students approach the task? What challenges might students face as they solve the task? **Planning to respond to student thinking:** What assessing questions might draw out student thinking? What advancing questions will help you move student thinking forward? **Planning to notice student thinking:** What strategies do you want to be on the lookout for as students work on the task? (Smith et al., 2020, p. 38)	**Key considerations** • Because variability is omnipresent in statistics, teachers must plan for many more possible responses/representations. • The teacher may not have data ahead of time in order to be able to anticipate exact responses (e.g., if data is being collected by students). • Anticipating how students might measure if collecting their own data. • Potential data analysis/visualization strategies based on data type (categorical vs. quantitative). • Potential views of data by students (case vs. aggregate perspectives). • Potential outcomes for variables/measures (if students collect the data). • Recognition of potential for data to deviate from anticipated distribution (if student collected). • Students will bring prior knowledge and experiences about the context you are exploring into their investigations and discussions. **Guiding questions for the teacher (during *planning*)** • What visualizations of data might be helpful to see different patterns or trends? • What visualizations of data might students rely on based on their view of data? (e.g., if they have a case view of the data? An aggregate view of the data?) What questions might advance students toward thinking about the data from an aggregate view? • What variables and analyses might students focus on for the particular data context? What conclusions might they draw? • Where might variation be present in student's investigations? (e.g., variation between samples, variation within a sample, sampling variability, measurement variability, etc.) How will you highlight this in your questioning? • How might data deviate from anticipated distribution (if student collected or simulated), and what will you do if the data do not align with anticipated patterns/trends? • What data cleaning/organizing issues might arise? • What topics might come up in discussing the context of the data? What resources might you want to have on hand to support students in understanding the context? • How do you answer students' questions relating to their own experiences with the context and outcomes from the data investigation?

students' work, but also how the students' work progresses and changes throughout the investigation. As teachers question students, they will also need to assess students' interpretations of the data and advance their thinking based on their learning goals: a process that can be complicated by the variety of statistical questions that students may have pursued in their investigations. In the selected teacher episode, we see how the teacher visits the groups and asks the students to share their exploration to monitor their work. As an example of an assessing question, the teacher asks a pair of students to explain an aspect of a graph they had made: "So, what do you think these different shades of green mean?" (Video 2, 1:36). The guidance of the teacher in this episode was also crucial to advance the thinking of a pair of students, encouraging them to consider two variables simultaneously (Video 2, 1:12). For some additional descriptions of the teacher's monitoring practices, see the supplementary material (link online). After monitoring student work, the teacher moves to selecting student work to highlight during the whole-class discussion (Practice 3; see Table 4).

Table 3 Key Considerations and Questions for Monitoring During a Data Investigation

Practice 2: Monitoring **Tracking student thinking:** How will you keep track of student responses and ensure that you check in with all students? **Assessing student thinking:** Are your assessing questions meeting students where they are and making thinking visible? **Advancing student thinking:** Are your advancing questions driven by your lesson goals? Are students able to pursue advancing questions on their own? Are your advancing questions helping students to progress? (Smith et al., 2020, p. 70)	**Key considerations** • Students may create several displays and/or do a variety of calculations to make sense of variation and patterns in the data as part of their investigation. • If using technology, their displays may change quickly as attributes are added or removed, making it a challenge to document a specific moment in the investigation. • Students may investigate different questions based on what they noticed in the data. • Students may investigate the data in ways which do not align with your learning goals in terms of the big statistical idea(s) you are focused on. • Students may not appropriately consider the context of the data in their analysis (e.g., drawing conclusions using the data which do not make sense based on the context or relying on their own experience with the context in their analysis rather than using the data to draw conclusions). • Students may not appropriately consider variability in their investigation and analysis. • Students may make statements implying certainty in instances where one cannot be certain. **Guiding questions for the teacher (during *investigation/group work*)** • If a group has multiple charts/visualizations, how will you document this while monitoring their work? • How will you document the trends/patterns that groups are noticing as you monitor their work? • How will you know if a particular group's approach is aligned with your learning goals? What features will you look for as evidence of this? • What will you say/do if groups verbally identify patterns or trends which are not supported by the data? • How is variability represented in their displays or statistics? What questions might you ask to highlight this? • How is the context of the data represented in their analysis and interpretation phases? • How do students create different arguments with the data and what evidence are they using to support those arguments? • What phase(s) of the investigation will you choose for them to present to the class and how will that promote students meeting the learning goal(s)?

Because of the fluid nature of data investigations, especially when students use technology like CODAP to explore data, selecting work to be presented may involve choosing not only the group/individual, but also what in their investigation is important for the class to discuss. Teachers may need to document specific representations in the moment—perhaps by taking a picture or screenshot—if they believe it is important to highlight later on, as it may be difficult to recreate the work once students have moved on in their investigations.

It is also essential that, when selecting work to be presented, teachers explore a variety of representations and consider how these may show similar or different features of the data set. For example, one representation may show a pattern over time, while another highlights distributional characteristics like center and variability. For instance, the scatterplot in Figure 3 allows one to investigate how the maximum height of roller coasters has changed over time, while the dotplot and boxplot in Figure 4 allow one to identify characteristics of the distribution of maximum heights of roller coasters more easily.

In the selected teaching episode, the chosen student pairs were selected because all of them investigated the relationship between height and speed of roller coasters, but each team chose a different approach and an associated type of graph to do so. This diversity in approaches allowed for a meaningful discussion focused on comparison and contrast of strategies by the selected pairs. The order in which students present their work is also important for a productive class discussion; this is the focus of Practice 4 (see Table 5).

Selecting and sequencing are closely connected when facilitating discussions about data contexts. Sequencing that acknowledges the different ways that students pursued their investigations is important because groups may have come to various conclusions based on which variables they considered in their displays and analyses. This is different from what occurs in typical mathematics lessons.

One potential pattern for sequencing focuses on movement from approaches that focus on the data as individual cases to those that take a more holistic approach (an aggregate perspective; Konold et al., 2015). In the selected classroom episode, the teacher carefully sequenced the team presentations according to this progression. The first student pair used two stacked dot plots to conclude that the fastest roller coaster also had the biggest drop (Video 2, 2:53). They were only focused on this one case in their analysis. The second pair of students created a dot plot of top speed color-coded by drop height (Video 2, 2:46); this

Table 4 Key Considerations and Questions for Selecting During a Data Investigation

Practice 3: Selecting **Identifying student work to highlight:** Which student solution strategies would help you accomplish your [statistical] goals for the lesson? What challenges did students face in solving the task? Were there any common challenges? **Purposefully selecting individual presenters:** Which students do you want to involve in presenting their work? How might selecting particular students promote equitable access to [statistics] learning in your classroom? (Smith et al., 2020, p. 104)	**Key considerations** • Student work will not necessarily include a “solution,” and instead will likely consist of multiple ways to use visualizations or statistics to support claims. • Students may have multiple visualizations to share, and they may not be able to easily “go back” to a specific display once they have moved on in their exploration. • Analyses done by different students may highlight different features of the data depending on which variables were investigated. • Students may have come to opposing, but reasonable, conclusions. • Emphasis upon data moves, like sorting, filtering, grouping, or cleaning the dataset can be important to share. • Decisions about whether sharing should include displays which uncovered similar features of the data, or those that differed in key ways. **Guiding questions for the teacher (during *planning of the discussion*)** • What features of the data do you want to be made evident during the discussion? • Which specific display or strategy from a group or individual do you want shared? Why? • Which visualizations show similar features of the data? Which show different features? • How do different representations work to answer the statistical question? • How will you navigate presentation of analyses that use the same data to draw different conclusions? • Which students need to be seen as competent?

supported them in looking more holistically at the relationship between top speed and drop height of roller coasters. The student pair purposefully placed last in the sequence of student presentations showed how they took an aggregate perspective to use a scatterplot to visualize the relationship between three variables simultaneously: top speed, drop height, and material type (Video 2, 6:19; see Figure 5). As

Figure 3 Graph Considering Maximum Height of Roller Coasters Over Time

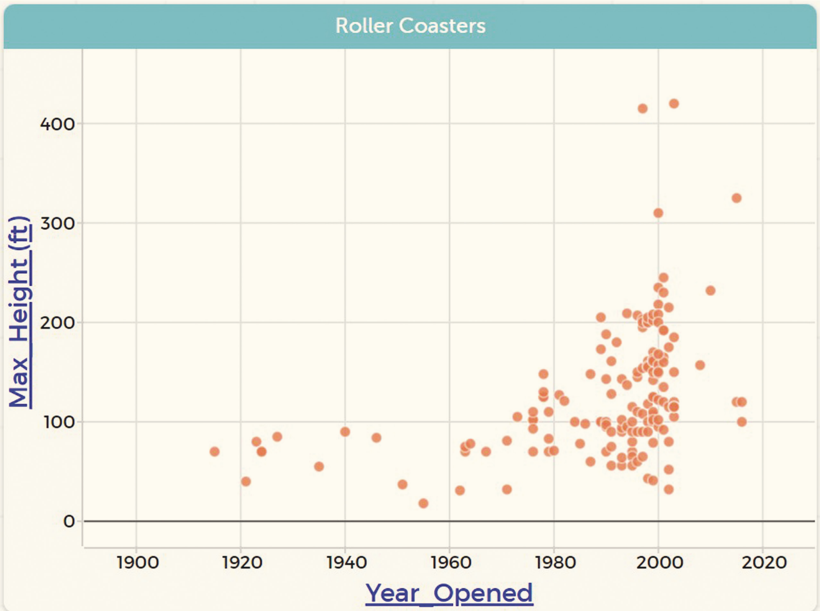
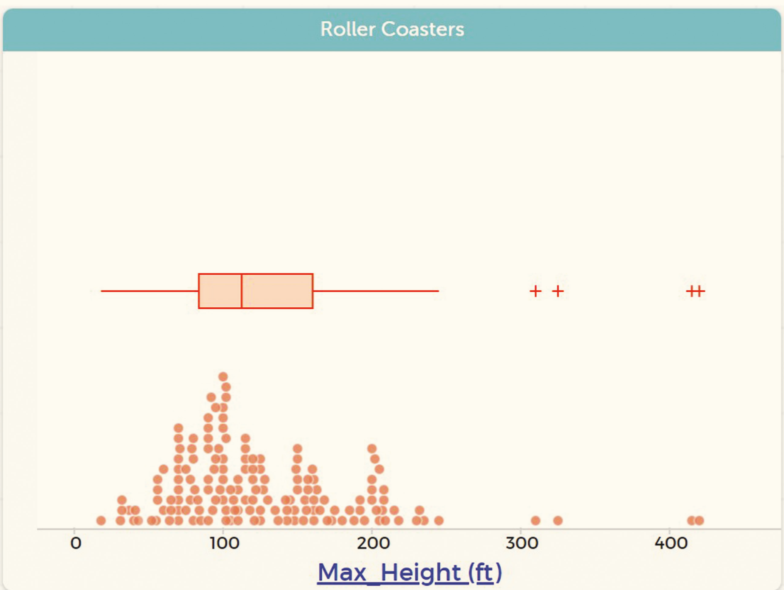


Figure 4 Graph Considering Distribution of Maximum Heights of Roller Coasters



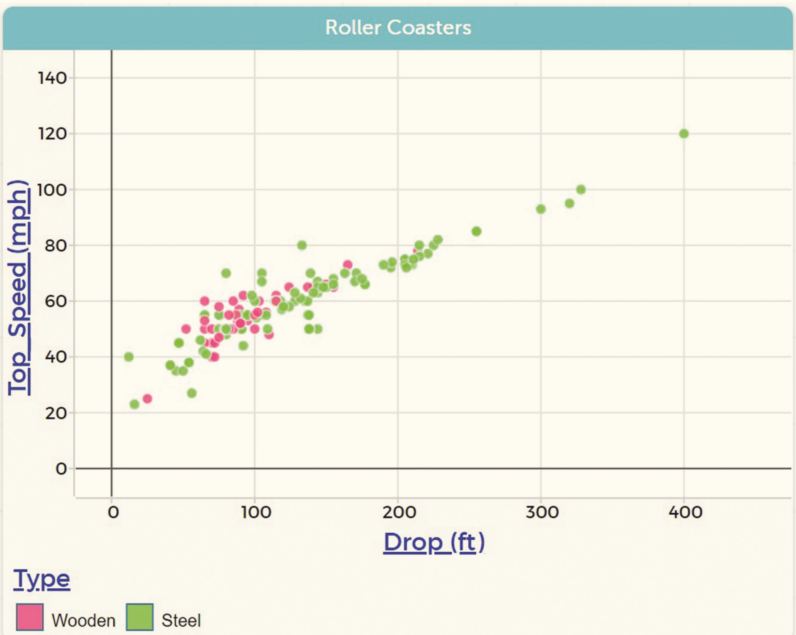
these students made their presentations to the class, the teacher highlighted important connections (Practice 5; see Table 6).

Making connections between students' approaches is critical during the class discussion when teaching mathematics or statistics. However, while these discussions around math tasks typically focus on varied strategies to reach a common solution, statistical data investigations are never 'solved' and the responses to the posed statistical investigation question may vary. Teachers need to not only connect students' responses to each other and their lesson goals, but also highlight which features of the data are more or less convincing in terms of the conclusions students have drawn. Comparing and contrasting students' displays and analyses is a key feature of this practice. In the selected classroom episode, the teacher contrasts students' data visualizations to make the differences between them explicit: "So, this

Table 5 Key Considerations and Questions for Sequencing During a Data Investigation

Practice 4: Sequencing Establishing a coherent storyline: How can you order the student work such that there is a coherent storyline related to the [statistical] learning goal? (Smith et al., 2020, p. 104)	**Key considerations** • May want to highlight multiple visualizations from the same individual or group. • Data cleaning/measurement/organization techniques may mean that students/groups were working with different datasets, resulting in differences in their investigations which may need to be considered in terms of order presented/discussed. • Students may have focused on different features of the data, making it difficult to create a "coherent storyline" that is accessible to all groups.
	Guiding questions for the teacher (during *planning of the discussion*) • Which data features do you want to highlight and in what order? • Which perspectives on the data should come earlier in the discussion? Which should come later? • How will aggregate reasoning and a focus on different kinds of variation be made evident during the discussion? • Which analyses highlight the patterns and trends in the data most clearly? • How can the data be interpreted in multiple ways?

Figure 5 Graph Considering Top Speed, Drop, and Type



graph kind of tells us some of the same information that the two graphs made but it's all in one graph" (Video 2, 5:10). Also, when asking if a later graph tells any new information, she is asking students to connect the various graphics the student pairs have presented. She was expecting students to see how the later graph supports viewing the data in aggregate (Video 2, 8:35). The supplementary material (link online) provides an additional example of this practice related to sampling.

CONCLUSION

While the example presented here was tailored for seventh-grade students, this framework can be used at all educational levels by adjusting responses to key questions and considerations. For instance, in Practice 0 at the elementary level, the choice of data set might be adjusted, replacing numerical data with categorical information, and incorporating technology might not be suitable. In Practice 1, younger learners may tend to focus on visualizations involving a case perspective, while older learners may gravitate more toward aggregation. During Practice 5, the expected level of formality in drawing inferences may vary based on the age of students involved.

The roller coaster activity can also be applied or adapted to elementary and high school levels, as the context is engaging for students of all ages. For example, at the elementary level, we could reduce the number of variables and the number of roller coasters in the table while providing students with more specific questions to investigate. In high school, we could focus on more formal statistical techniques, such as confidence intervals, the line of best fit, or residuals.

We offer the adapted 5 Practices to support teachers in using a common framework to enact more powerful discussions during statistical investigations. The GAISE II document states that a goal of statistics education in PK–12 classrooms is for students to "be data problem solvers who interrogate the data and use questioning throughout the statistical problem-solving process to make decisions with confidence, understanding that the art of communication with data is essential" (Bargagliotti et al., 2020, p. 8). As the quotation emphasizes, we can view statistical problem solving and communication as both an art and a science. Students must learn to consider multiple ways of organizing and representing data to answer a statistical question. Unlike in many mathematics tasks, there is not one right way to solve or respond to a statistical question. Therefore, teachers can use this adapted 5 Practices framework to prepare for the many possibilities within any statistical investigation and provide students with meaningful opportunities to make sense of the world through data.

Table 6 Key Considerations and Questions for Connecting During a Data Investigation

Practice 5: Connecting **Connecting student work to the goals of the lesson:** What questions about the student work make the [statistics] being targeted in the lesson visible? **Connecting different solutions to each other:** What questions will help students make connections between different solution strategies presented? (Smith et al., 2020, p. 142)	**Key considerations** • Drawing inferences in a statistical investigation is different from coming to a "solution" to a task in mathematics. Students must consider not only their conclusions, but how confident/certain they are in claims about a population based on their sample data. • Connections must be made between various data visualizations and/or statistical measures (e.g., highlighting how various data visualizations can bring out different features of the data). • An additional connection teachers need to focus on in discussions of statistical investigations is to the context of the investigation. **Guiding questions for the teacher (during *leading of the discussion*)** • How are the different ways that we are visualizing the data functioning to answer the statistical question? • How is variability represented in different visualizations? • How are different perspectives on data present? (i.e., aggregate vs. case views of data) How might you connect these during the discussion? • What does the analysis of this sample data tell (or not tell) about the population? How will you emphasize this in the discussion? • Given the context of the investigation, are there limitations of any of the visualizations to be noted?

REFERENCES

Bargagliotti, A., Franklin, C., Arnold, P., Johnson, S., Perez, L., & Spangler, S. (2020). *Pre-K–12 guidelines for assessment and instruction in statistics education II (GAISE II): A framework for statistics and data science education*. American Statistical Association.

Byun, S., Weiland, T., Cannon, S., Fernandes, A., Nti-Asante, E., Peterson, F., Smucker, K., & Engledowl, C. (2023). Teaching and learning with data investigation: Working group report from 2022. In T. Lamberg & D. Moss (Eds.), *Proceedings of the forty-fifth annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 674–678). University of Nevada Reno.

Cobb, G. W., & Moore, D. S. (1997). Mathematics, statistics, and teaching. *The American Mathematical Monthly*, *104*(9), 801–823. https://doi.org/10.1080/00029890.1997.11990723

Concord Consortium. (2023). *Common online data analysis platform (CODAP) [computer software]*. https://codap.concord.org/

Franklin, C. A., Bargagliotti, A. E., Case, C. A., Kader, G. D., Scheaffer, R. L., &, Spangler, D. A. (2015). *Statistical education of teachers*. American Statistical Association. https://www.amstat.org/asa/files/pdfs/EDU-SET.pdf

Groth, R. E. (2015). Using the five practices model to promote statistical discourse. *Teaching Statistics*, *37*(1), 13–17. https://doi.org/10.1111/test.12052

Konold, C., Higgins, T., Russell, S. J., & Khalil, K. (2015). Data seen through different lenses. *Educational Studies in Mathematics*, *88*(3), 305–325. https://doi.org/10.1007/s10649-013-9529-8

Makar, K., & Rubin, A. (2009). A framework for thinking about informal statistical inference. *Statistics Education Research Journal*, *8*(1), 82–105.

Mojica, G. F., Azmy, C. N., & Lee, H. S. (2019). Exploring data with CODAP. *Mathematics Teacher*, *112*(6), 473–476. https://doi.org/10.5951/mathteacher.112.6.0473.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*.

National Council of Teachers of Mathematics. (2014). *Principles to action: Ensuring mathematical success for all*.

Smith, M., Bill, V., & Sherin, M. G. (2020). *The 5 practices in practice: Successfully orchestrating mathematics discussions in your elementary classroom*. Corwin Press.

Smith, M., & Stein, M. K. (2018). *5 practices for orchestrating productive mathematics discussions*. National Council of Teachers of Mathematics.

Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, *67*(3), 223–248.

ACKNOWLEDGMENTS

The authors would like to thank those involved in the ESTEEM project for creating the video resources that were used to showcase how the proposed framework might look in action. We would also like to acknowledge the Teaching and Learning with Data Investigations working group at the 44th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education, where the idea and impetus for this work emerged (Byun et al., 2023).

This paper is partially supported by the National Science Foundation under Grants DUE 2141727 and 1625713 awarded to North Carolina State University, DUE 2141716 awarded to Eastern Michigan University, and DUE 2141724 awarded to University of Southern Indiana. Any opinions, findings, and conclusions or recommendations expressed herein are those of the authors and do not necessarily reflect the views of the National Science Foundation.

5 practices at a glance (what each practice might look/sound like during a statistical investigation)

<u>**Planning**</u>: **In preparing for the classroom, this might look/sound like . . .** ● Considering different exploration tools and their affordances (i.e. CODAP, Fathom, Excel, SAS DataFly, creating graphs by hand), in order to decide which makes sense for your learning objectives and access to technology. ● Choosing a context where students collect data themselves when appropriate to increase engagement and student consideration of study design and data collection issues. ● Picking a data set that makes sense for your context (for example, selecting a small data set with under 20 values for an elementary setting).
<u>**Anticipating**</u>**: In preparing for the classroom, this might look/sound like . . .** ● Creating a variety of visualizations to prepare for potential student approaches. ● Noting questions you might ask groups or individuals based on each potential visualization's features. ● Having "backup" additional data to add to student collected results when data do not align with anticipated patterns (usually due to a low sample size). ● Having conversations about the issue you are exploring and what you are learning from the data about that issue.
<u>**Monitoring**</u>**: In the classroom, this might look/sound like . . .** **Example Assessing Questions:** ● Can you tell me what each _____ represents? ● What does this tell us about the distribution of the data? ● What does this tell us in regards to the question we are investigating? ● How was this data collected and how does that impact what we can say from it? **Example Advancing Questions:** ● Do you think these patterns would hold for a different population? ● How might the analysis change if we added additional data points or additional variables? ● How might the analysis change if the study design had been different (e.g. experimental vs observational)? ● What statistical question could you ask with the data you have?
<u>**Selecting**</u>**: In the classroom, this might look/sound like . . .** ● Taking a screenshot of a group's work at a specific stage in their investigation in order to highlight it later during the discussion. ● Based on what you anticipated from the data investigation, selecting work which highlights the different features you identified (for example, different forms of variation, different relationships between variables, etc.).
<u>**Sequencing**</u>**: In the classroom, this might look/sound like . . .** ● Determining which displays are showcased/shared on their own, and which should be displayed at the same time. ● Acknowledging at the beginning of the discussion that groups may have come to varying conclusions based on the data they chose to focus on (especially if a data set with many variables was used).

<u>Connecting</u>: In the classroom, this might look/sound like . . .

- Asking students to gauge or quantify their confidence/certainty in conclusions drawn by themselves and/or their classmates.
- Planning to showcase two specific visualizations side-by-side to compare and contrast their features.

More detail on the 5 Practices in the Roller Coaster task videos

Five practices	Enactment of the practice in the teaching episode
Planning	● The teacher obtained and used a dataset about roller coasters in the U.S. that consisted of 157 roller coasters in 32 states. ● The teacher investigated key and strategic facts about context, for example, about the history and materials of roller coasters. ● The teacher found a video that would launch the lesson and highlight attributes of roller coasters that could kickstart the students' investigations. ● The teacher thought about an initial question to start the investigation (What is the maximum height that wooden roller coasters tend to be?) that would be of interest to students.
Anticipating	● The teacher anticipated attributes in the roller coaster dataset that students were likely to be interested in (e.g., type of material and maximum height) to purposefully bring those out in an engaging way (Video 1). ● Students can arrive at the same conclusions using different methods, especially when using dynamic software such as CODAP. In the video, one pair of students linked top speed with length simply by selecting the last point on the speed graph and observing that it coincided with the last point on the length graph (Video 2, 0:49). In contrast, another team used two different graphs displayed simultaneously to reach the same conclusion (Video 2, 3:24).
Monitoring	● Teacher visits the groups and asks the students to share their explorations. The teacher asks for explanations for specific characteristics or trends in data/graphs on the screen. T: "*So, what do you think these different shades of green mean?*" (Video 2, 1:36) ● Teacher gives some advice about how to improve the current visualization in CODAP (Video 2, 1:05). In the process, the pair of students learn a new feature of CODAP that they could use in another visualization. ● Students may generate inferences or narratives that are not immediately apparent from the data, which can be valuable. However, teachers should strive to steer the conversation back to the information that the

	data provide in order to foster statistical reasoning, while also acknowledging and incorporating students' viewpoints. This can be achieved by, for example, rephrasing the same or a similar question. T: *So, there's steel in them you know. Alright, what can we tell from this graph about wooden and steel coasters? (Video 2, 9:06)* ● Teacher encourages another pair of students to recreate an interesting graph they had made (Video 2, 2:02). We see that the monitoring stage is also a time for the teacher to assess student progress, pose advancing questions, give feedback on their work thus far, and extend their thinking. ● Initially, a pair of students analyzed the roller coaster dataset case-by-case by color-coding dots on a graph, focusing only on the fastest roller coaster. However, with the guidance of their teacher, they were encouraged to consider two variables simultaneously (maximum height and top speed) in an aggregate manner. (Video 2, 1:12)
Selecting	● The student pairs were selected because all of them investigated the relationship between height and speed of roller coasters, but in different ways. This diversity in approaches allowed for effective comparison and contrast of strategies in the resulting whole class discussion.
Sequencing	● The teacher carefully sequenced the student pairs' presentations, starting with a case-by-case approach and gradually progressing to an aggregate approach. This can be observed in the video (Video 2, 2:53), where the teacher guides the students through a progression from one variable to three variables. The first team used stacked dot plots to focus on the fastest roller coaster and noticed that it also had the biggest drop. The second team created a top-speed dot plot colored by drop height, and the third team used a scatterplot to visualize the relationship between top speed, drop height, and material type.

Connecting	● The teacher contrasts the approaches of the different student pairs. T: *So, this graph kind of tells us some of the same information that the two graphs made but it's all in one graph.* (Video 2, 5:10) ● The teacher promotes reflection about the relationship between the sample and the population. T: *Do you think those are all the roller coasters from the U.S.?* (Video 2, 6:32) ● When asking if a later graph tells any new information, she is asking students to to connect the various graphics the student pairs have presented. She was expecting students to see how the later graph supports viewing the data in aggregate. (Video 2, 8:35)