Boise State University
ScholarWorks

Mathematics Faculty Publications and Presentations

Department of Mathematics

5-1-2018

Secondary Mathematics Teachers' Planned Approaches For Teaching Standard Deviation

Maryann E. Huey
Drake University

Joe Champion
Boise State University

Stephanie Casey
Eastern Michigan University

Nicholas H. Wasserman
Columbia University

SECONDARY MATHEMATICS TEACHERS' PLANNED APPROACHES FOR TEACHING STANDARD DEVIATION

MARYANN E. HUEY
Drake University
maryann.huey@drake.edu

JOE CHAMPION
Boise State University
joechampion@boisestate.edu

STEPHANIE CASEY
Eastern Michigan University
scasey1@emich.edu

NICHOLAS H. WASSERMAN
Teachers College, Columbia University
wasserman@tc.columbia.edu

ABSTRACT

Research-based guidelines for learning variation exist (e.g., Franklin et al., 2007; Garfield, delMas, & Chance, 2007), but little is known about how teachers plan to teach standard deviation, or how these plans align with recent recommendations. In this article, we survey lesson plans designed by in-service and preservice secondary mathematical teachers. We report on the accuracy, technology usage, and visual representations in the lesson plans. We consider how many elements are used, the level of conceptual development, and the mathematical nature. Findings support differences between preservice and master's level students in education, as well as a tendency by in-service teachers to teach in alignment with prior learning experiences, despite professional development. Implications for teacher education and curricular development are offered.

Keywords: *Statistics education research; Teacher preparation*

1. INTRODUCTION

In many countries throughout the world, statistics has recently been included as a core component of primary and secondary mathematics programs (Senior Secondary Board of South Australia, 2002; USA–Common Core State Standards Initiative [CCSSI], 2010). Variation, often referred to as the heart of statistics (Watson, Kelly, Callingham, & Shaughnessy, 2003), is a foundational concept essential to the learning and understanding of statistics at all levels of schooling and an international learning goal in many secondary mathematics programs (e.g., Brazil–Ministério da Educação, 2006; South Africa–Department of Education, 2002; USA-CCSI, 2010). The *Guidelines for Assessment and Instruction in Statistics* (*GAISE*) *Report*: *A Pre-K-12 Curriculum Framework* suggests a progression in relation to learning about variation (Franklin et al., 2007). Building upon prior experiences with measures and conceptions of variation, students learn about standard deviation, the most common and conceptually challenging univariate measure of variation, in secondary school. Although the Pre-K-12 *GAISE* recommendations are evident in the articulation and sequencing of the *Common Core State Standards for*

Statistics Education Research Journal, 17(1), 61-84, http://www.stat.auckland.ac.nz/serj

Mathematics (*CCSS-M*) adopted throughout the United States (CCSSI, 2010), few secondary teachers learned statistics in a similar manner. The vast majority of in-service secondary teachers likely encountered standard deviation in tertiary coursework with an emphasis placed upon procedures and computation (Garfield & Ben-Zvi, 2007).

Hence an immediate issue exists—teachers' prior experiences and preparation are known to be inadequate in relation to current secondary mathematics learning goals. While programmatic recommendations are evolving, such as *The Statistical Education of Teachers* guidelines (Franklin et al., 2015), little is known about how teachers plan to teach statistics topics, including standard deviation, and whether such approaches align with research-based instructional recommendations. Given relatively new teacher preparation materials and guidance, such as Garfield and Ben-Zvi's (2008) Statistical Reasoning Learning Environment model, uncertainty exists regarding how preparation incorporating these ideas translates to actual instruction. Long-term questions exist in relation to how to prepare teachers to teach this relatively new progression of statistical ideas included in many secondary mathematics programs throughout the world.

In this paper, we report upon how preservice and in-service secondary mathematics teachers with varying levels of prior knowledge, teaching experiences, and exposures to current pedagogical recommendations plan to teach standard deviation. We focus upon the instruction of standard deviation due to the emphasis placed on this variation measure during secondary and tertiary statistics coursework. Whereas many statistics topics are new in nature to the secondary mathematics curriculum, standard deviation has been included in curricular materials for decades. However, teachers find this topic especially challenging and tend to view standard deviation as cumbersome to teach (Reading, 2004). Given the interconnected web of statistics concepts that comprise an understanding of standard deviation, a focus upon planned instruction of standard deviation provides a unique view into teachers' knowledge of variation and their associated instructional choices. This research responds to the urgent need for investigating teachers' knowledge of variation and their associated instructional practices (Sánchez, Silva, & Coutinho, 2011) in order to inform statistics and mathematics educators and curriculum developers.

2. BACKGROUND AND RESEARCH QUESTIONS

Standard deviation is known to be difficult for students to conceptualize and compute, creating unique challenges for instruction (Reading & Shaughnessy, 2004). In this paper, we adopt delMas and Liu's (2005, 2007) notion that in order to develop a robust conceptual understanding of standard deviation, students need multiple opportunities to productively coordinate the foundational concepts of distribution, mean, and deviation of data from the mean. With the current curricular sequencing of statistics topics by grade in the *CCSS-M*, the foundational concepts of distribution, variation, and center are introduced in middle school, as well as two measures of variation: mean-absolute deviation (MAD) and interquartile range (IQR) (CCSSI, 2010; Franklin et al., 2007). Secondary mathematics teachers in the United States are guided to plan instruction that builds upon the written curriculum, as specified by standards, policies, recommendations, and textbooks. Teachers' assumptions of what students have experienced previously and what they should learn next are largely based upon the curricular standards. However, teachers' learning experiences are likely to be very different from the progression of topics outlined in the *CCSS-M*, as well as recommendations put forth in documents such as the Pre-K-12 *GAISE* framework as "most instruction on standard deviation tends to emphasize teaching a formula, practice with performing calculations, and tying the standard deviation to the empirical rule of the normal distribution" (delMas & Liu, 2005, p. 56). In the following review of literature, we first synthesize research related to instruction that supports students' understanding of measures of variation via a hypothetical learning trajectory, and then summarize prior research specific to teachers' understanding of standard deviation. Lastly, we summarize curriculum research in order to describe how teachers' lesson plans can inform our understanding of instruction and the related opportunities to learn about standard deviation provided to students.

2.1. STUDENTS' UNDERSTANDING AND LEARNING ABOUT VARIATION MEASURES

In an effort to improve instruction on the topic of variation and variation measures, Garfield, delMas, and Chance (2007) embarked upon a teaching experiment utilizing Japanese Lesson Study. The experiment involved iteratively assessing tertiary students' understandings of variation in a college-level statistics course and then designing research-based lesson plans to address findings. In addition to producing new lesson plans, the team created a hypothetical learning trajectory for variation that embodied both their new findings as well as prior research findings. We have adapted this trajectory, rephrasing their stages of statistical reasoning in terms of key understandings to be addressed by instructional activities; doing so allowed us to capture pivotal learning opportunities in teachers' lesson plans for developing a robust understanding of variation with a specific focus on standard deviation (Figure 1). We employ the definition of statistical reasoning offered by Garfield and Ben-Zvi (2004), "as the way people reason with statistical ideas" and "connect one concept to another" (p. 7). Further, statistical "reasoning means understanding and being able to explain statistical processes" (p. 7). The adapted trajectory includes stages of understandings and associated learning experiences that require connecting, coordinating, and reasoning about core statistical concepts, ideas, and processes.

We utilized the adapted trajectory as a framework for our work because it embodies a plethora of research findings and recommendations regarding instructional experiences that benefit students' learning from the beginning stages of understanding variation through the final stage of comparing and contrasting measures of variation. Although the trajectory is depicted as sequential, the organization illustrates a progression of advancing understandings that could be learned in a variety of orders and combinations and need not focus on only one level of understanding. In fact, the students in Garfield et al.'s (2007) study had prior exposure to many of the included topics. Next, we briefly describe how their hypothetical learning trajectory was developed and highlight relevant research.

Prior to the teaching experiment, Garfield et al. (2007) recognized that understanding variability measures requires both informal and formal aspects, and that transitioning from informal conceptions of variation, that is, data values are different, to understanding, reasoning about, and interpreting formal measures of variation, such as standard deviation, is particularly challenging and difficult. During the study, Garfield et al. found that tertiary students needed learning opportunities at the most basic level or stage, labeled as 1 on Figure 1, related to how and why data vary. From assessment results, students struggled to see variation in data when represented graphically in histograms. Garfield et al. also felt strongly that students' intuitive conceptions of variation and ways of reasoning about variation must be considered and serve as the basis for all new learning. In a new lesson developed by Garfield et al., students were provided with categories of data based on survey results from their own class. Students were asked to predict which categories would have high or low variation, provide justifications for their choices, and then sketch graphical representations of the data. The researchers purposely utilized data that were relevant to their students in order to support their thinking about variability. In alignment with McClain, McGatha, and Hodge's (2002) recommendations, students were then provided with an opportunity to explain their thinking and methods in an effort to deepen understandings. Afterwards, students compared their hypothetical graphs to the actual data distributions and were provided with statistical measures of variation. This lesson spanned the first two levels of understandings on Figure 1 labeled stages 1 (understanding that data vary) and 2 (understand how variation can be visualized). Post-assessment results related to this lesson demonstrated growth in students' understandings of what variation meant and students were able to identify data distributions with more or less variation using graphical displays. A descriptor of the key instructional activities is provided on the right side of Figure 1 for both stages.

In stage 3, the understanding is "variation can occur in clusters of data." The need for this understanding arose based on students' observed proclivity to describe variation in terms of the range and shape of data distributions without reference to central tendency or clusters of data. Some students also highlighted the presence of outliers as increasing the variation again without reference to the distance of

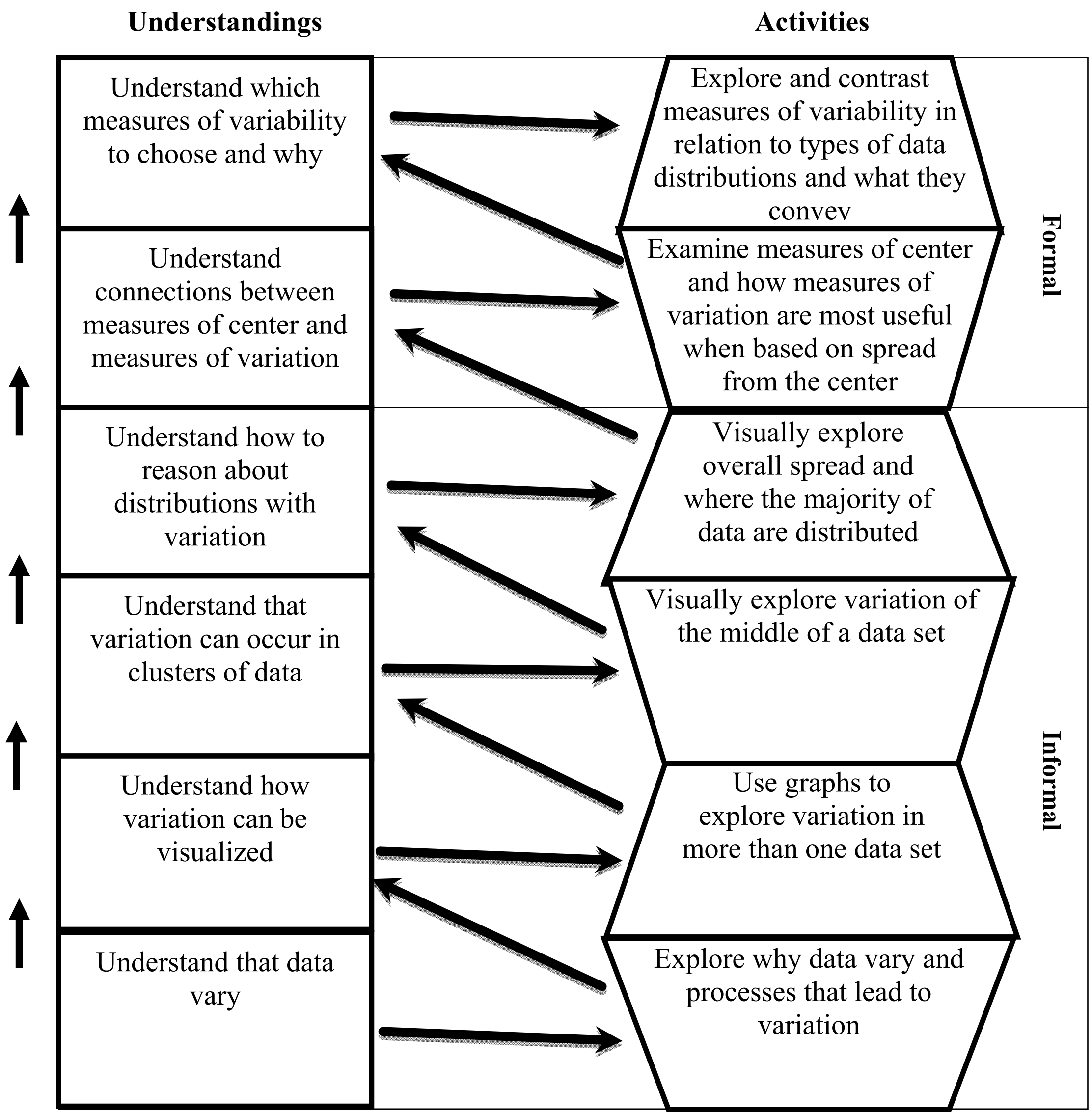


Figure 1. Instructional activities for developing students' key understandings related to variation measures. Adapted from "Using students' informal notions of variability to develop an understanding of formal measures of variability," by J. Garfield, R. delMas, & B. Chance, 2007, in M. C. Lovett and P. Shah (Eds.), *Thinking with data*, pp. 117–147, New York: Erlbaum.

data values away from a central tendency measure or cluster of data. Other research studies have found similar trends in students' descriptions of variation in data distributions (delMas & Liu, 2005). Therefore, specific instruction and attention is needed in this regard. However, these recommendations were made after completion of the teaching experiment. In a related research study, Lehrer, Kim, and Schauble (2007) engaged grade 5 and 6 students in an extensive data modeling process. The students measured various real-world objects and arrived at differing values through measurement error. Students sought ways to quantify the precision of their measurements dynamically by analyzing data distributions with the aid of technology. Through this process, students invented their own measures for variation, viewed in this case as measurement error relative to main clusters of data and measures of central tendency.

Garfield et al. (2007) also acknowledged that strong distributional understandings are essential to learning about variation, described as an awareness of the overall spread and where the majority of data are distributed, which is the understanding in stage 4. However, details on activities were not provided. There are numerous studies related to supporting students' development of distributional reasoning (e.g., Bakker & Gravemeijer, 2004). The assumption at this particular stage is that formal measures of variation, such as standard deviation and IQR, need not be incorporated into the learning experience. Rather, the understanding relates to seeing data distributions as an entity, including trends in the data as well as spread. In this way, students are coordinating two statistical ideas and reasoning about the set of data holistically.

In a short assessment following the first lesson, students were asked how their understandings of variability were impacted after seeing summary statistics, such as IQR and standard deviation (Garfield et al., 2007). They were also asked about what criteria they would use to gauge high or low variability in the future. Similar to other research studies, students did not seem to have an understanding of the need for or utility of variation measures beyond range, although they had previously learned about both IQR and standard deviation. Rather, their responses tended to focus upon the shape of the data, center, outliers, skew, range, or combinations of these ideas. A few responses included references to standard deviation, but none referred to IQR. Watson and Moritz (2000) suggest that students must be presented with learning experiences that necessitate representing variation of data with a single summary value. Mentioned previously, Lehrer et al. (2007) engaged grade 5 and 6 students in an extensive data modeling process that developed a need for a single measure of variation in sets of data.

Garfield et al. (2007) designed one additional new lesson for students involving comparing and contrasting measures of variation graphically and numerically in order to reason about which data set had the most variation. Students were provided with multiple data sets and the associated measures of range, IQR, and standard deviation. The numerical values provided conflicting rankings, so students had to use and reason about all relevant aspects of the data to arrive at an ordering. This lesson targeted understandings in stages 5 and 6 in Figure 1, as students had to determine which measure of variation best represented the data based on the appropriate measure of central tendency and then interpret what this measure meant in relation to the data distributions.

In a similar effort involving tertiary students, delMas and Liu (2005) created a series of histogram-based tasks to advance students' understandings of how the magnitude of standard deviation measures relate to the distance of data values away from mean. The sequence of activities increased students' reference to the distance of data values away from the mean in their justifications, but only a few students discussed how the frequency of values at various distances influenced the magnitude of the standard deviation measure. It is important to state that whereas Garfield et al. (2007) observed improvements in students' understanding of variation, including standard deviation, the authors expressed that additional learning experiences were needed.

Our adaptations of the hypothetical learning trajectory are relatively minimal. We identified and separated the learning activities and associated understandings at each stage of the trajectory (Garfield et al., 2007). Stages of understanding are labeled 1 through 6 in Figure 1 for ease of reference. We have added the distinction of informal and formal to the framework to denote where formal measures of center and variation are introduced and explored.

2.2. SECONDARY TEACHERS' KNOWLEDGE OF MEASURES OF VARIATION

Though less literature has addressed teachers' understandings of variability, the few existing studies suggest many teachers' understandings are similar to those of students, with the possible exception of teachers who are regularly responsible for teaching college-level statistics courses in secondary school settings (Peters, 2014). For example, Hammerman and Rubin (2004) noted that secondary teachers involved in professional development discussed variation using only segments and slices of data distributions, rather than viewing variation as a characteristic of the entire distribution similar to Bakker and Gravemeijer (2004). Confrey and Makar (2002) also found that middle school teachers tended to examine variation in data distributions by focusing on individual points instead of global characteristics similar to Garfield et al.'s (2007) findings. When Makar and Confrey (2005) analyzed how 17 preservice teachers talked about variation in data distributions, they reported a wide variety of informal interpretations, such as clustered, clumped, grouped, etc., but very little use of standard statistical language such as mean, standard deviation, and skew. Whereas most of the teachers were comfortable in utilizing standard deviation as a measure to characterize variation, they had difficulty expressing in words what the measure meant in relation to the data.

In another study, Silva and Coutinho (2008) reported a spectrum of in-service middle and secondary mathematics teachers' conceptions of standard deviation. Some teachers described standard deviation as a measure of dispersion of data, but were unable to understand how it might be employed to characterize a set of data. Others realized that a higher standard deviation meant more spread in the data and a lower value meant less variation, but either identified the central referent as the mode or did not refer to one at all. Several teachers also felt that a lower standard deviation was preferable in some way in comparison to a higher standard deviation, but could not verbalize why this was the case. The most sophisticated analyses of data distributions consisted of coordinating range with a central cluster of data or modal data, but did not include mean or standard deviation, two topics that had been discussed previously. Other teachers reported standard deviation in their analyses, but did not explain what it meant in relation to the data. Predominantly, teachers depicted standard deviation as a measure of homogeneity of the data without reference to a measure of central tendency. Based on these findings, we were especially interested to see how teachers would plan to present standard deviation to secondary students.

In summary, ample research suggests that both students and teachers struggle to state what standard deviation represents. Research findings suggest that students and teachers need opportunities to connect and develop their informal notions of core statistical ideas prior to being taught formal definitions and procedures through carefully sequenced instructional experiences (Garfield et al., 2007; Makar & Confrey, 2005). Promising interventions share common characteristics, such as the utilization of graphical displays of data distributions, often dynamically represented with technology, in order to allow students and teachers to manipulate displays, make sense of data and measures, and draw conclusions about relationships between measures of center and variation (delMas & Liu, 2007). In addition, productive tasks often encourage exploration of variation in data sets without prescribed procedures and problem contexts that highlight the need for a summary statistic of variation (Lehrer et al., 2007; Shaughnessy, 2007). Next, we discuss the role of curriculum in students' opportunities to learn.

2.3. CURRICULAR STUDY AND LESSON PLAN ANALYSIS

The written curriculum, as defined by policies, recommendations, and textbooks, plays an integral role in the design of mathematics instruction. Documents such as the *Common Core State Standards for Mathematics* (CCSSI, 2010) and the Pre-K-12 *GAISE* recommendations (Franklin et al., 2007) provide learning goals in relation to variation and are part of the written curriculum. Before students have an opportunity to learn from the written curriculum, whether a textbook or another resource, the curriculum unfolds and changes through a series of temporal transformations (Stein, Remillard, & Smith, 2007). These phases are mediated by a variety of factors, such as teachers' knowledge and beliefs, orientation towards curriculum, classroom structures and norms, and organizational and policy contexts. Following the written curriculum, the second temporal phase consists of the intended curriculum, which is the teacher's plan for instruction or lesson plan. The third phase is the enacted curriculum and refers to what actually takes place in the classroom and denotes the implementation phase leading directly to student learning or the attained curriculum.

Given the relatively recent adoption of the Common Core State Standards for Mathematics in the United States (CCSSI, 2010) and similar adoptions around the world in relation to statistics, the examination of teachers' intended curriculum has never been more important, particularly in content domains in which teachers have little experience (CCSSI, 2010). By studying a variety of teachers' lesson plans, we will have a view into teachers' prior understandings, beliefs, and goals in relation to standard deviation, as well as evidence of what they think will lead to student learning within an actual classroom. These findings will help inform the design of teacher preparation programs and professional development efforts.

Beyond providing guidance to statistics and mathematics teacher educators, this analysis can inform the degree to which teachers are likely to provide students with meaningful opportunities to learn about standard deviation in secondary schools. Despite most textbooks containing several chapters dedicated to statistics and probability, Tarr, Chávez, Reys, and Reys (2006) report that secondary mathematics teachers tend to omit statistics and probability lesson plans at higher rates compared to other content areas. Further, researchers observed that although standards emphasize the importance of teaching statistics, changes in actual classrooms tend to lag behind recommendations (Tarr & Jones, 2010). One possible explanation proposed is that teachers have not experienced learning statistics content and processes in alignment with current standards.

The purpose of this research study is to investigate secondary mathematics teachers' planned instruction for teaching standard deviation. Specifically, the research questions addressed by this study are

1. What characterizes secondary mathematics teachers' lesson plans for teaching standard deviation?
2. Do teachers' lesson plans differ based on their prior coursework, teaching experiences, and exposure to pedagogical recommendations?

By answering these questions, we create a view of teachers' intended curriculum for standard deviation and consider the influence of several factors.

3. METHODOLOGY

3.1. SETTINGS AND PARTICIPANTS

As an exploratory study, one of our primary goals was to contribute to a broad understanding of teachers' intended curriculum for teaching standard deviation across prior statistics coursework, teaching experience levels, and exposure to current pedagogical recommendations specific to statistics. In order to reveal differences across these factors, we recruited both preservice and in-service teachers as participants and implemented a simultaneous research protocol in four settings: (a) a sophomore preservice statistics content course designed for future mathematics teachers in the northwestern part of the United States, (b) a graduate mathematics course at a large private university in the northeast part of the United States required for a master's degree program in education, (c) a senior undergraduate level mathematics teaching methods course at a large public university in the middle of the United States, and (d) an in-service teacher professional development program for statistics and probability teaching in the middle portion of the United States. The research team purposefully recruited a total of 16 participants with the three goals of spanning a spectrum of content knowledge based on prior coursework, balancing the participation of preservice and in-service teachers, and balancing exposure to current statistics education recommendations. Within each institution, participants were sought who attended class on a regular basis, attempted assignments in a conscientious manner, and who, we believed, were interested and willing to engage in a reflective process regarding their lesson plan assignments. Participants were compensated with $15 gift cards.

Table 1 summarizes participants' teaching experience, academic standing, prior coursework and education specific to statistics, as well as, probability content, and exposure to current statistics education teaching recommendations, such as the Pre-K-12 *GAISE* recommendations (Franklin et al., 2007) and the high school statistics and probability learning progressions (The Common Core Standards Writing Team, 2012). All participants were assigned pseudonyms for privacy.

Before describing the participants' courses and professional development, it is important to note that this study was conducted in the last quarter of each class or professional development program, and the

Table 1. Summary of participants' teaching experience, prior coursework, and exposure to statistics teaching recommendations

Participating teacher (Institution)	Preservice/ In-service (Years of teaching)	Program Type	Prior college statistics and probability coursework	Exposure to recommendations for teaching statistics
Shawna (A)	Preservice	Undergraduate	Statistical Methods	Limited
Bill (A)	Preservice	Undergraduate	Statistical Decision Making in Business	Limited
Tonia (A)	Preservice	Undergraduate	None	Limited
Alexa (B)	Preservice	Graduate (Undergraduate: Mathematics)	Probability Theory, Mathematical Statistics	Moderate
Tom (B)	In-service (0)	Graduate (Undergraduate: Mathematics Education)	Introductory Probability and Statistics	Moderate
Elise (B)	In-service (3)	Graduate (Undergraduate: Mathematics Education)	Introductory Probability and Statistics	Moderate
Debbie (C)	Preservice	Undergraduate	Calculus-based statistics	Moderate
Ted (C)	Preservice	Undergraduate	Introduction to Statistics, Calculus-based statistics	Moderate
Brian (C)	Preservice	Undergraduate	Calculus-based statistics	Moderate
Jay (C)	Preservice	Undergraduate	Calculus-based statistics	Moderate
Ellen (C)	Preservice	Undergraduate	Calculus-based statistics	Moderate
Maria (D)	In-service (7)	Professional Development	Introductory Probability and Statistics	Extensive
Dylan (D)	In-service (4)	Professional Development	Introductory Probability and Statistics	Extensive
Anna (D)	In-service (16)	Professional Development	Introductory Probability and Statistics	Extensive
Feliz (D)	In-service (18)	Professional Development	Introductory Probability and Statistics	Extensive
Lacy (D)	Inservice (20)	Professional Development	Introductory Probability and Statistics	Extensive

Note. Graduate or Undergraduate: Participants' undergraduate degrees are shown in parentheses. Exposure to recommendations for teaching statistics: The degree of exposure to current recommendations is relative to others in the study based on prior coursework focused on statistics teaching methods.

lesson planning assignment was embedded within these learning experiences as a requirement. The participants from institution A were all preservice teachers in their second year of college. They were enrolled in a statistics content course designed for secondary mathematics education majors, which included an assigned reading of the Pre-K-12 *GAISE* recommendations and framework during the first weeks of the course (CCSSI, 2010). Students then applied the recommendations in course assignments throughout the semester. The textbook utilized was *Statistics: Informed Decisions Using Data, 3rd Edition* by Sullivan (2010), and students engaged in activity-based learning throughout the semester including use of TI calculators, GeoGebra, applets, and Excel. The institution provided early teaching experiences for students throughout their first and second years of coursework. Therefore, the students had prior experiences writing and teaching several mathematics lesson plans with an emphasis placed on inquiry-based instruction, which was the pedagogical focus for K-12 mathematics and science education at this particular institution. Two of the participants completed college-level statistics coursework previously, Statistical Methods in Psychology and Statistical Techniques for Making Business Decisions, and the third participant did not have prior coursework in statistics or probability.

The participants from institution B were all graduate students pursuing master's degrees in mathematics education. One of the students completed an undergraduate mathematics degree, and the other two completed mathematics education degrees. Only one of the participants taught high school mathematics prior to the study, as the others transitioned to graduate school immediately after completing their undergraduate degrees. The students were enrolled in a topics in theoretical probability course designed specifically for graduate students in mathematics education. The required books for the course were *Introduction to Probability* by Kelly (1994) and *The Drunkard's Walk: How Randomness Rules Our Lives* by Mlodinow (2008). The class focused on advanced probability topics through a mathematical lens with connections to secondary mathematics teaching. As an undergraduate student, one participant completed two separate courses for probability and statistics while pursuing a mathematics degree. The other two students previously completed an introduction to probability and statistics class.

The five participants from institution C were all preservice teachers in their last year of undergraduate college. They were enrolled in their only mathematics methods course designed to develop content-specific pedagogical knowledge. All participants had previously completed a calculus-based statistics course and one had also completed an introductory course for statistics. The textbooks utilized were *5 Practices for Orchestrating Productive Mathematics Discussions* by Smith and Stein (2011) and *Teaching Mathematics in Grades 6-12: Developing Research-Based Instructional Practices* by Groth (2013). Assignments and readings included understanding and applying the Pre-K-12 *GAISE* recommendations and framework (Franklin et al., 2007), as well as, the *CCSS-M* (CCSSI, 2010). All participants had extensive classroom experiences related to mathematics prior to the study; however, a practicum or classroom-based component was not included with this particular methods course.

Finally, the participants from institution D were all rural, in-service teachers of high school mathematics. Three teachers had from 16 to 20 years of teaching experience, and two had 7 and 4 years of teaching experience. One of the teachers, Lacy, actively taught a college-level statistics course, as part of her regular teaching responsibilities. Another, Feliz, had prior experiences teaching the topics of correlation and linear regression to Algebra 2 students. The others did not regularly teach probability or statistics content, but would soon be required to integrate this content into traditional algebra and geometry courses given the adoption of the content standards outlined in the *CCSS-M* (CCSSI, 2010). All had completed an introduction to probability and statistics content course during their undergraduate preparation, and these teachers were near the end of a year-long professional development program focused on preparing high school teachers to teach statistics and probability in alignment with current recommendations. Teachers primarily learned probability and statistics content during two weeks in the summer by engaging in activity-based lessons in collaborative settings. The activities spanned middle and secondary level content with extensions into college level content. The teachers learned how to use TinkerPlots™, Fathom™, applets, Core Math Tools, and their TI calculators during the professional development program, and they developed their own lesson plans in professional learning communities for the academic school year. They were provided with a number of curricular resources from the National Council of Teachers of Mathematics, the American Statistical Association, the Pre-K-12 *GAISE* framework (Franklin et al., 2010), and *Workshop Statistics: Discovery with Data* (Rossman & Chance, 2011).

3.2. DATA SOURCES

The data sources for this study were electronic submissions for a lesson planning task, complimented by participants' verbal explanations of key features and decisions related to lesson plans during semi-structured interviews. The lesson planning task prompted participants to "write a lesson plan for which the objective is to have students understand standard deviation" with attention to three related content standards in the *CCSS-M* (CCSSI, 2010). See the Appendix for the full prompt. Participants were asked to describe in detail the key ideas, tasks, and activities that would be included to ensure students gained a solid understanding of standard deviation.

In this introductory lesson plan, we did not outline what understanding standard deviation meant to participants. However, we did provide fairly direct guidance that instruction should build upon students' prior learnings of mean, distribution, MAD and IQR. We also emphasized that the lesson plans should attend to key ideas and understandings explicitly. Therefore, our expectations were that participants would develop students' conceptual understanding of what standard deviation measures, the typical distance of data values away from the mean, and compare and contrast this in some way with the measures of MAD and IQR either in terms of the computations or ideas. This corresponds to stage 5 of the framework and leads into stage 6 (see Figure 1). In stage 5, students learn to examine measures of center and understand how measures of variation are most useful when based on spread from the center. With the guidance provided to participants, they were to assume that stage 5 learning had previously occurred in relation to IQR and MAD during prior grades. The key understanding in stage 5 is that measures of spread are more valuable when considered with a reference to a measure of center, highlighting the connections between these two ideas.

Once students understand conceptually the information that standard deviation and other measures of variation convey about data distributions in reference to measures of center, they compare and contrast variation measures in stage 6 (see Figure 1). Learning experiences for students at this stage should allow them to notice characteristics of variation measures in relation to types of data distributions. Through these comparisons, students gain an understanding of the utility of each variation measure in relation to the data distribution at hand.

We provided three standards from the *CCSS-M* to serve as different ways to think about the role of standard deviation and the associated learning goals (CCSSI, 2010). The rest of the assignment was open-ended in nature, so that participants felt unconstrained regarding their approach to the lesson planning assignment. Our intent with this design was to mimic the challenge that many teachers currently face. They are asked to teach to standards without curricular materials readily available or identified. With this design, we attempted to elicit what teachers would actually plan for this particular topic of standard deviation.

Finally, semi-structured interviews regarding the lesson plan submissions were conducted, and investigators interviewed participants from institutions that they were not employed by for 15 to 20 minutes in duration. Interviewers asked participants to explain their rationale for the lesson plan overall followed by an explanation of each portion of the lesson plan. Interviewers inquired about how participants anticipated students engaging with the activities and what they would learn. In addition, choices related to technology, data representations, and context were briefly discussed. Finally, if lesson plans included a formula, we would inquire about the choice of formula and how they planned to describe it to their students, when applicable. The audio portion of the interviews were recorded and later transcribed. During the interviews, participants were provided with a paper copy of their lesson plan and were asked to talk through their lesson with the interviewer, highlighting the important ideas and tasks while explaining their goals and rationale.

3.3. ANALYSIS

The four researchers divided into teams of two to code the lesson plan submissions and interview transcripts. The coding of interview transcripts supplemented the axial coding of the lesson plan submissions. Phrases and exchanges that captured instructional choices, resources utilized, concerns or areas requiring revisions, misconceptions, and beliefs about what experiences students need to learn standard deviation were captured and included as notes with each participants' axial codes. The initial axial codes used to analyze the lesson plans originated from prior research findings and recommendations for teaching standard deviation and variation more broadly. These codes were refined and modified in an iterative manner in order to capture key aspects and themes found in lesson plan submissions when raters met to compare codings. Codes were also added as needed, which we will describe later. All lesson plan submissions were independently double-coded, and we maintained a codebook with clear operational definitions and descriptors (Miles & Huberman, 1994). Once we reached agreement on the codings for the subset of lesson plans, we continued with an iterative process of independently double-coding subsets of responses and meeting frequently to compare, discuss areas of disagreement, and resolve any discrepancies.

Axial codes for analyzing lesson plan submissions informed the research questions: i) What characterizes secondary mathematics teachers' lesson plans for teaching standard deviation? and ii) Do teachers' lesson plans differ based on their prior coursework, teaching experiences, and exposure to pedagogical recommendations? Initial codes were created from learning activities depicted in Figure 1 and described in Section 2.1. The codes were aligned to Stages 5 and Below or Stage 6 based upon the understandings likely developed in the learning activities. Stages 5 and Below codes relate to developing understandings of standard deviation as a measure of variation. General activities related to understanding variation more broadly would certainly fall into this category as well; however, teachers did not include such activities given the prompt to write a lesson plan specific to standard deviation. Stage 6 codes focus upon understanding which measure of variability to choose and why (see Figure 1). Recall, a lesson plan could certainly contain codes from multiple stages, similar to lesson plans developed by Garfield et al. (2007). For consistency, we analyzed the first day of any multi-day lesson plans submitted by participants. We also coded the working definition of standard deviation employed in each lesson plan and noted errors or omissions, such as not referencing distance away from the mean. Garfield et al. (2007) and delMas and Lui (2005, 2007) convey the critical need for students to understand that standard deviation represents a typical distance of data values away from the mean. This aspect is captured in the coding related to the standard deviation definition.

In Table 2, our coding scheme is shown. Items italicized relate directly to productive learning activities identified by Garfield et al. (2007) and other researchers. Additional codes were added for unanticipated responses, and original codes were modified as needed to align more closely with responses provided. For example, a Stage 5 and Below sub-item code for activities involving the empirical rule was added later. A Stage 6 code was added based upon one lesson plan motivating the use of standard deviation through comparisons to range. Although this was not a recommended instructional choice by Garfield et al. (2007), we included this code in order to capture the intentions of this participant. The types of visual representations (e.g., line plot, histogram, etc.) were coded to gauge whether students were encouraged to visually reason about variation in data distributions (delMas & Liu, 2005, 2007), and finally, we also coded for type of technology within plans, a common trait of successful learning experiences (Lehrer et al., 2007; delMas & Liu, 2005, 2007).

4. RESULTS

The results of the study demonstrate a wide variety of planned approaches to teaching standard deviation. In this section, the lesson plan submissions are discussed by institutional cohort followed by an analysis across cohorts to examine trends based on teaching experiences, prior knowledge, and exposure to current statistical teaching recommendations.

4.1. LESSON PLANS

Institution A The participants from this institution were enrolled in a statistics content course for future teachers and in their second year of college. In comparison to the other cohorts, these participants represent the lowest level of prior knowledge as defined by prior statistics coursework, least amount of teaching experience, and limited exposure to current pedagogical recommendations specific to teaching statistics again based on prior coursework. However, the participants had engaged in designing and delivering inquiry-based lessons in mathematics in other courses.

Table 2. Coded elements and sub-item codings for lesson planning submissions

Included elements within the lesson plans (yes or no).
- Sub-items coded within present elements

Definition of standard deviation: A full definition of standard deviation is developed.
- Errors in the definition.
- Omission in the definition, such as the distance away from the mean.

Stages 5 and Below: Developing understanding of variation and standard deviation as the typical distance of data away from the mean.
- *Interpret what standard deviation means or represents in the context of a problem.*
- *Visually analyze variation within or between data sets represented graphically or numerically.*
- *Compare standard deviation values of two data sets to determine which has more variation.*
- *Analyze the influence of outliers on standard deviation values.*
- Apply the empirical rule with normal distributions to identify typical and unusual values.
- Generate an inference about population data or significant differences between two data sets using increments of standard deviation.

Stage 6: Understand which measures of variability to choose and why
- *Comparison of MAD and/or IQR to standard deviation for the same set of data either visually or numerically.*
- *Exploring when to use standard deviation versus MAD and/or IQR.*
- Compare standard deviation to range as a more accurate or informative measure of variation.

Visual representations: Visual representations are employed either by the teacher or created by students.

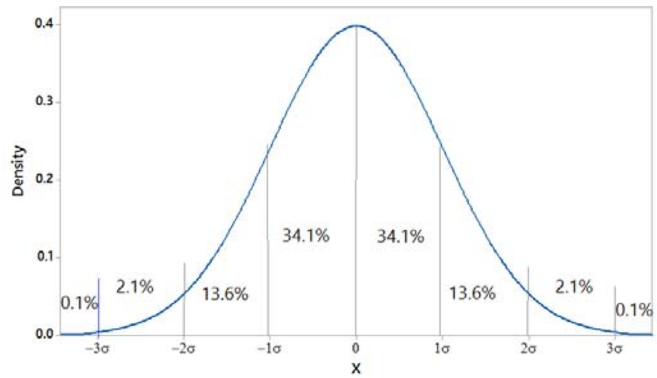
Technology: Technology is utilized with the lesson plan.

Note. Italicized sub-items indicate activities that are included in Figure 1 as productive learning experiences. All types of visual representations and technology tools were captured for each participant's lesson plan.

Each teacher's lesson plan included one or two Stages 5 or Below productive activities, including interpretations of standard deviation values or comparison of standard deviation values between data sets. Tonia had students compute and compare standard deviation measures of actual quarterback data to see who was the most consistent player. However, she did not require students to fully describe what standard deviation values represented in this context. Instead, she simply asked which quarterback was most consistent based on measures. The two other lesson plans from this cohort required students to interpret what computed standard deviation measures represented in relation to the data provided, usually set in real-world contexts.

These same two lesson plans included other activities involving the empirical rule, an added sub-item for Stages 5 and Below. One plan contained activities involving generating an inference, another added sub-item for Stages 5 and Below. In her lesson plan, Shawna asked students where most values lay by identifying the interval of one standard deviation away from the mean, an application of the empirical rule. Another participant, Bill, asked students which values were unlikely or more than three standard deviations away from the mean using the empirical rule. He also asked how standard deviation measures could be applied in a business context (see Figure 3) and included tasks where students analyzed overlapping regions of standard deviation measures to gauge whether significant differences in mean values of two sets of data existed, an activity related to generating inferences, an added sub-item in Stages 5 and Below.

Discuss what σ = 29.6 means in terms of the data. Show students the bell curve and define the confidence intervals as on the graph. Discuss why confidence intervals might be important. For example, a clothing store might like to know the average height for adult men in the United States is about 70 inches with a standard deviation of around 3 inches. How would that standard deviation be helpful to Levi Strauss Inc.? Another question: Could σ = 0?



Note: We left the graph as originally portrayed within the lesson plan recognizing that the labeling is not completely accurate unless the mean value is zero.

Figure 3. One of Bill's empirical rule tasks involving typical heights of males

Notably, Bill was the only participant from this cohort who attempted to connect to previously learned measures of variation. He asked students to calculate the range, IQR, and standard deviation for a data set and say which one was "best" without guidance on how to make this decision, a sub-item in Stage 6. Bill was also the only participant in this cohort to not include a full definition of standard deviation in the lesson plan materials. He wrote that standard deviation is "how far apart all the data in a set are." Although he asked students to interpret the meaning of standard deviation, evidence is not provided that students would learn that standard deviation represents the average distance of data values from the mean.

In relation to use of graphical displays, the two lesson plans emphasizing the empirical rule included normal distributions similar to what is shown in Figure 3. One histogram was provided in Bill's lesson plan during the introduction to show how many hours of television people watch per day. Tonia's lesson plan of quarterback data did not include any use of visuals, but relied entirely upon numerical measures. In terms of technology, two lesson plans utilized Excel, but only as a way to make computations more efficient. The other employed calculators for computing the standard deviation. All participants had students work through the step-by-step process of computing standard deviation versus leveraging available functions within the calculators or spreadsheet applications.

In summary, this cohort provided lesson plans involving the normal distribution more than other cohorts likely due to the content they were learning in class. Each lesson plan either included interpreting standard deviation measures or comparing standard deviation measures, which are productive activities depicted within our framework for Stages 5 and Below. One participant included connections to previously learned measures of variation, but omitted the role of the mean in the definition of standard deviation.

Institution B The participants from this institution were enrolled in an advanced topics in probability theory course for graduate students in a mathematics education master's program. Alexa had completed a mathematics undergraduate degree previously, and Tom and Elise had completed undergraduate degrees in mathematics education. Of the three, Elise was the only one with prior teaching experience as a licensed teacher, but Tom had completed his student teaching assignment and was a licensed teacher. These participants represent a high degree of prior coursework related to statistics, moderate teaching experiences, and limited exposure to pedagogical recommendations about teaching statistics, as the recommendations were not emphasized during this particular class nor during undergraduate coursework. They are similar to beginning teachers, but presumably have more content knowledge.

In terms of understanding standard deviation as a measure of variation, all three lesson plans asked students to interpret what standard deviation measures in relation to sets of data, a productive learning activity aligned with Stages 5 and Below of our framework. Tom designed a lesson where students recorded how high they could jump. Then, students calculated the standard deviation to illustrate the variation of jump heights for their class and interpret what this measure represented. Tom did not connect to previously learned measures of variation, and the only graphical display was a number line for marking the jump height of students.

Alexa crafted a discovery-based lesson plan that encouraged students to discover what standard deviation measures, coded as interpreting standard deviation in Stages 5 and Below, by using calculator functions to compute, without formulas, the mean, median, mode, range, IQR, and standard deviation of multiple data sets. Students were asked to compare the standard deviation values of various data sets to the associated IQRs, a Stage 6 activity, and guess what standard deviation measures. Alexa's lesson plan included a variety of graphical displays such as histograms, line plots, and stem-and-leaf plots.

The third lesson by Elise engaged students in a variety of activities and ways of thinking about standard deviation. Initially students computed the standard deviation for multiple sets of data and interpreted what the values represented, a Stages 5 and Below activity. Then students were asked to choose members of a running relay team based on measures of center and variation by comparing standard deviation values, another Stages 5 and Below activity. The last activity, again Stages 5 and Below, asked students to visually determine which data distribution had a larger standard deviation given symmetric line plots of data with the same mean value. At the end of the lesson plan, Elise positioned standard deviation as a superior or more accurate measure of variation in comparison to MAD through the two data sets: $\{-4, -4, 4, 4\}$ and $\{7, 1, -6, -2\}$, which have the same MAD measurement, but not the same standard deviation. Through this brief illustration, Elise compares standard deviation and MAD numerically, a Stage 6 sub-item, which we coded as comparing standard deviation to MAD numerically.

All participants from this institution provided and requested full descriptions of standard deviation from students including a reference to the mean. The use of technology was similar to that of Institution A. Students used either handheld calculators or Excel spreadsheets to calculate standard deviation step-by-step with the exception of Alexa's lesson plan. She had students use built in functions on calculators to compute all measures. Graphical displays of data were abundant in one lesson plan, but only minimally utilized in others. Tom included only one number line of data, and Elise included two line plots in one task.

In summary, this cohort provided lesson plans that all attended to interpreting the meaning of standard deviation measures, a Stages 5 and Below activity, and provided full definitions of standard deviation. In addition, two lesson plans made comparisons to previously learned measures of variation, Stage 6 activities. This cohort was also the only group to have activities that all aligned with recommendations originally contained in our framework.

Institution C Participants from Institution C were enrolled in a mathematics teaching methods course and near the end of their college coursework. These participants would enter their student-teaching experience in the following semester and are representative of what we would expect of beginning teachers in the United States, who have fulfilled course requirements but not yet begun to teach full-time. Overall, these participants completed a medium degree of prior coursework related to statistics and probability, accrued moderate teaching experiences, and were moderately exposed to pedagogical recommendations about teaching statistics in relation to other groups within our study.

Two sets of lesson plans were similar for this cohort and one was unique. Debbie and Ted created lesson plans involving data collection followed by graphing, computing standard deviation, and interpreting what standard deviation measures, a productive learning activity aligned with Stages 5 and Below of our framework. One additional component of Ted's lesson plan involved the identification of outliers using the empirical rule, that is, values more than three standard deviations away from the mean are considered outliers, which was a sub-item not originally contained in our framework, but added later. Neither lesson plan included references to prior measures of variation beyond computation.

For the second set of lesson plans, Brian and Ellen's lesson plans focused on studying the effect of outliers on standard deviation and visually analyzing variation, two sub-items of Stages 5 and Below. In addition, both lesson plans made connections to IQR, Stage 6 sub-items, but in different ways. Brian asked students to compute standard deviation and IQR for the same data sets, compare the values, a Stage 6 sub-item, and explain why the measures were similar or different given the presence of outliers. Then, students graphed sets of data in order to predict how standard deviation values would change if outliers were removed. Ellen also had students compute standard deviation and IQR and create line plots of the data sets. Students compared the IQR to the interval of one standard deviation away from the mean to see which measure made more sense depending upon the existence of outliers in the data. In this way, Ellen had students explore when to use standard deviation versus IQR based on characteristics of the data distribution, a Stage 6 sub-item.

Jay provided a lesson plan that did not have any Stages 5 and Below elements, but rather positioned standard deviation as a superior measure of variation in comparison to MAD using sets of size four in the same way as Elise, which are easily accessible via an internet search. We coded this comparison as a Stage 6 sub-item. Jay was the only participant in this cohort to not fully state the definition of standard deviation, and said that it "described the spread of the data" in his lesson plan. Jay's lesson plan was notable in that he did not include any visual displays of data that would support an informal analysis of variation, and he did not include any reference to the mean.

In terms of visual displays of data, both Debbie and Ted left the choice of graphical displays open but required students to create one. Brian and Ellen both employed primarily line plots, and Brian had one normal distribution illustrating the percentages of data contained within standard deviation increments away from the mean. Jay did not include any visual representations. Technology played a role similar to other cohorts. Three teachers used a combination of functions for computations and step-by-step procedural calculations, and two included only procedural calculations with the aid of Excel or handheld calculator technology.

In summary, this cohort provided lesson plans involving multiple Stages 5 and Below sub-items, most of which aligned with the original framework adapted from Garfield et al. (2007): interpreting what standard deviation represents in three cases, visualizing variation in two cases, studying the effects of outliers on standard deviation measures in another case, and applying the empirical rule. In addition, three lesson plans made connections to previously learned measures of spread (Stage 6 sub-items) included in the original framework. This cohort generated the most codings in comparison to other cohorts. One lesson plan was not as robust and failed to include the full definition of standard deviation.

Institution D Participants from Institution D were in-service teachers completing a year-long professional development program that included an array of technology tools, curricular materials, and current pedagogical recommendations for teaching statistics. This cohort represents a moderate degree of prior coursework, as the professional development focused upon secondary school content as well as pedagogy, and a high degree of both prior teaching experiences and exposure to current recommendations in comparison to other cohorts in this study.

Two sets of lesson plans were similar for this cohort and one was unique. Dylan and Lacy both attended to Stages 5 and Below sub-items of visually analyzing variation between data sets represented graphically and numerically and comparing standard deviation values of two data sets to see which has more variation. Dylan had students collect data on the frequency of various colors of M&M candies in a package. Students computed the mean and standard deviation for each color, created line plots of the data, and then assessed whether the standard deviation values matched the variation observed in the dot plots. In addition, Dylan asked students to compare the standard deviation measures numerically between colors. As a final task, Dylan asked students to rank the variation of data sets without making computations by visually analyzing the numeric values of the data (see Figure 4).

Each of the five sets of numbers all have a mean of 100. Rank the data sets from 1-5, with 1 representing the set of data with the lowest standard deviation WITHOUT calculating the standard deviation.

_____a.	90	95	100	105	110
_____b.	25	100	200	0	175
_____c.	100	100	100	100	100
_____d.	100	80	120	75	125
_____e.	300	0	0	100	100

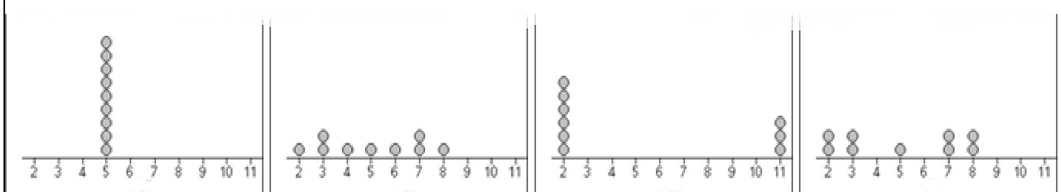
Give a full, detailed explanation of why you chose the ranking that you did.

Figure 4. Dylan's variation task involving a visual analysis of numeric values

Similarly, Lacy also asked students upon several occasions to visually compare sets of data and graphical displays of data to determine which varies the most without computations, a Stages 5 and Below sub-item. In addition, Lacy asked students to interpret what standard deviation represented in the context of a problem, another Stages 5 and Below sub-item (see Figure 5).

1. Draw a dot plot that has a relatively high standard deviation.
2. Draw a dot plot that has a relatively low standard deviation.
3. Rank these data sets in order from 1 to 4 (without a calculator). The set with the least standard deviation should be ranked the number 1; the set with the greatest standard deviation should be ranked the number 4. [Note: All of the distributions pictured have a mean of 5.]

Rank ____ Rank ____ Rank ____ Rank ____



4. A sample of students at your school were asked how many hours of homework did they complete last week. Calculate and interpret the standard deviation in context for this data.
5, 6, 2, 8, 7

Figure 5. Lacy's visual analysis of variation and interpretation of standard deviation tasks

Both Dylan and Lacy employed the use of line plots in their lesson plans and calculator functions to compute standard deviation. Lacy also had students complete a step-by-step computation one time. Lastly, Dylan did not offer a definition for standard deviation beyond saying that "the larger the spread, the larger the deviation," nor did he reference any other previously learned measures of variation. Lacy did provide the full definition, and she briefly mentioned previously learned measures of variation including MAD and IQR as being less sophisticated.

Feliz motivated her lesson plan through the need for manufacturing companies to control the quality of their products by minimizing defects or outliers. Students collected data on the number of candies in packages and later on hand span measurements, calculated the standard deviations, and then identified outliers using the empirical rule, an added sub-item of Stages 5 and Below. No visuals were employed and connections to previously learned measures of variation included only a brief mention to MAD in terms of computational similarities with standard deviation. Feliz did include a complete definition of standard deviation.

Anna and Maria provided lesson plans that did not have any Stages 5 and Below elements, but rather asked students to compute standard deviation values. Anna provided two sets of data that had the same range and mean values, but different standard deviation measures to motivate the need for a more precise measure, we coded this as a new Stage 6 sub-item involving comparisons of range to standard deviation. Maria simply asked students to compute the standard deviation using the sample and population formulas. Maria and Anna gave accurate definitions for what standard deviation measures, but neither included visual displays.

In summary, two lesson plans focused upon productive activities aligned with Stages 5 and Below of the framework, Dylan's and Lacy's. Feliz's lesson plan connected to identification of outliers using the empirical rule, which was an added Stages 5 and Below sub-item. The other two lesson plans were procedural in nature. Notably, only one of the lessons asked students to interpret a standard deviation measure in the context of the data provided, which was unusual in comparison to other institutions. The role of technology was similar to other cohorts with a focus on handheld calculator computations and functions.

Summary of Stages 5 and Below sub-items in lesson plans Recall, the italicized items in Table 3 represent productive activities described by Garfield and colleagues' (2007) hypothetical learning trajectory for variation as well as other researcher's findings (see Figure 1). The other items were added from participants' lesson plans. In Table 3, we summarize the presence of codes for each sub-item in relation to lesson plans. Some lesson plans included activities that spanned multiple sub-items. Only one lesson plan did not include any sub-items related to Stages 5 and Below.

Table 3. Lesson plan summary of Stages 5 and Below sub-items

Developing conceptions of variation and standard deviation as the typical distance of data away from the mean.	Number of plans (%)
Interpret what standard deviation means or represents in the context of a problem.	9 (56%)
Visually analyze variation within or between data sets represented graphically or numerically.	5 (31%)
Compare standard deviation values of two data sets to determine which has more variation.	4 (25%)
Analyze the influence of outliers on standard deviation values.	2 (13%)
Apply the empirical rule with normal distributions to identify typical and unusual values.	4 (25%)
Generate an inference about population data or significant differences between two data sets using increments of standard deviation.	1 (6%)

The majority of lesson plans included an emphasis on interpreting the meaning of standard deviation, but not all. The types of learning activities included in lesson plans demonstrate variety more than consistency in relation to developing an understanding of standard deviation as a measure of variation.

Summary of Stage 6 sub-items in lesson plans In Table 4, we show the distribution of codes for each sub-item in relation to lesson plans that included Stage 6 sub-items. Four lesson plans did not make any reference to previously learned measures of variation, and five made only a brief mention of previously learned measures.

Predominantly, participants did not make meaningful connections or comparisons to previously learned measures of variation in lesson plans. Of those who did, the emphasis was placed on numeric comparisons and deciding when to use standard deviation versus IQR or MAD.

Table 4. Lesson plan summary of Stage 6 sub-items

Understand which measures of variability to choose and why	Number of Plans (%)
Comparison of MAD and/or IQR to standard deviation for the same set of data either visually or numerically.	4 (25%)
Exploring when to use standard deviation versus MAD and/or IQR.	2 (13%)
Compare standard deviation to range as a more accurate or informative measure of variation.	1 (6%)

4.2. RELATIONSHIP BETWEEN LESSON PLAN ELEMENTS, PRIOR COURSEWORK, TEACHING EXPERIENCE, AND EXPOSURE TO PEDAGOGICAL RECOMMENDATIONS

In this section we share how participants' lesson plans differed based on their prior coursework, teaching experiences, and exposure to pedagogical recommendations summarized in Table 1. Table 5 provides a summary of findings aggregated by institutional cohort.

Table 5. Summary of lesson plan elements by institutional cohort

Institution	At least one conceptual element (Stages 5 and Below sub-items)	At least one graphical visualization included	Connections to IQR or MAD beyond mentioning (Stage 6 sub-items)	Full definition of standard deviation
A	100% (100%)	67%	33% (33%)	67%
B	100% (100%)	67%	67% (67%)	100%
C	80% (60%)	80%	60% (60%)	60%
D	100% (60%)	40%	20% (0%)	80%

Prior Coursework Relationships The cohort with the highest level of prior statistics and probability coursework, Cohort B, tended to provide the complete definition for standard deviation within lesson plans. All lessons included an accurate and complete definition. Recall that all participants from Cohort B had completed required undergraduate degree coursework in statistics and probability and were enrolled in a graduate level course including content related to probability.

Teaching Experience The cohort with the most teaching experience, Cohort D, tended to incorporate visuals less often than other cohorts, with 40% of lesson plans compared to 67% and 80% of other cohorts. In addition, the Cohort D did not make meaningful connections to previously learned measures of variance with only one lesson plan, 20%, including a Stage 6 sub-item. This cohort also had the lowest levels of overall alignment with productive learning activities originally included within the framework based on Garfield et al. (2007) and other researchers, with 60% of lesson plans including Stages 5 or Below sub-items and 0% including Stage 6 sub-items.

Exposure to Pedagogical Recommendations Cohorts B and C had moderate exposure to pedagogical recommendations specific to statistics and tended to have the highest levels of making meaningful connections to previously learned measures of variation (Stage 6 sub-items), at 67% and 60% respectively.

5. DISCUSSION OF FINDINGS

5.1. CHALLENGE OF CHANGING PRACTICE

We were disappointed that Cohort D, the cohort with the highest level of teaching experience, did not include more productive learning activities identified by Garfield et al. (2007) and other researchers, as many of these activities are present in current pedagogical recommendations. The results of Cohort D, comprised entirely of in-service teachers who had recent experiences learning about standard deviation using technology with dynamic displays of data distributions, illustrate how challenging it is to break the habit of traditional ways of teaching. We were further surprised that none of these in-service teachers developed lesson plans that meaningfully connected to previously learned measures of variation beyond a brief mention. These measures were focused upon in detail during their professional development and emphasized in the lesson planning prompt. One possible explanation for these results is that in-service teachers generally do not write new lesson plans for a topic, but rather find an activity or lesson that is of interest and then adapt it for their classes. It is unclear if these selections would be any closer to the productive learning activities supported by research, as many available activities and lesson plans are not peer-reviewed or vetted. In addition, only one of the five in-service teachers included an activity that required students to interpret what standard deviation means or represents in the context of a problem, which would seem to be the basis for any initial lesson plan on this topic.

Encouragingly, the master's level teachers in mathematics education (Cohort B), two of which were new teachers, tended to write lesson plans that strongly aligned with both Stages 5 and Below learning activities as well as Stage 6 activities as depicted by Garfield et al. (2007) and other researchers. Coupled with their accurate definitions of standard deviation, this particular cohort was able to identify meaningful activities for learning that were statistically correct. These teachers strove to develop the meaning of standard deviation through productive experiences more consistently than other cohorts. Cohort C, which represents beginning teachers, had the highest number of productive learning activities overall, but struggled at times with the definition of standard deviation. It is important to note that Cohort C participants had experienced a calculus-based statistics course, which one might think would predispose them to mathematical approaches to statistics and a continuation of traditional ways of teaching. This certainly seemed to be the case based on Cohort A's tendency to lean towards application of the empirical rule and a focus on the normal distribution, as this was likely a focus of the course in which they were enrolled. However, both beginning teachers and recently graduated teachers in this study demonstrated value for recent pedagogical recommendations and were able to incorporate them into lesson plans.

5.2. STRUGGLE WITH WHY AND HOW TO TEACH STANDARD DEVIATION

Based on the lesson plans in this study, participants struggled with the purpose of an introductory lesson plan on standard deviation. One participant positioned standard deviation as the most sophisticated of all measures of variation. Two others positioned standard deviation as superior to MAD. A challenge exists to debunk the idea that statistical knowledge learned later in the curriculum is somehow superior to prior learning, as is often presumed to be the case in mathematics. Given the findings from Garfield et al. (2007) that tertiary students benefited from exploring informally their intuitive ideas about variation, this likely is the case for secondary students as well. By positioning standard deviation as a more sophisticated measure, the emphasis of the learning becomes more formal in nature versus exploratory and intuitive. Statements such as these also discount students' prior knowledge of variation, sending a hidden message that prior experiences and informal ways of thinking are not encouraged or relevant, which will impede students' development of the relationships between core statistical concepts and measures of variation. Many participants did attempt to connect to students' prior learning regarding measures of variation and situated learning activities as a natural progression, but appeared to struggle with how to do so productively. Many mathematics teachers feel compelled to answer the silent question in high school classrooms, “Why should I learn this?,” which can be a difficult question to answer when the teacher is unsure of the potential learning progression or how all the pieces might fit together for a given topic.

Moreover, the diversity of approaches in lesson plans suggests a lack of widespread awareness among mathematics teachers about how to effectively teach standard deviation. The sheer diversity of instructional approaches and foci that were present in these sixteen lesson plans (e.g., including or not including the calculation of standard deviation, the role of technology, how to motivate the need for discussing standard deviation, working with larger data sets or very small ones, plots of data or lists of data, normal curves or any distribution, etc.) suggests a haphazard statistical lesson-planning process. Important conceptual understandings about variation and how students might develop them seemed to be unknown to or unaccounted for by participants. Clearly, increased efforts are needed in teacher education programs specific to the teaching and learning of standard deviation. In addition, further research investigations should consider addressing the extent to which the findings in this report align with existing curricula, and possibly compare and contrast the variety of plans we observed in instructional approaches for standard deviation to what might be observed in lessons developed by mathematics teachers for other secondary statistics topics. Lastly, research is needed to study the attained curriculum specific to standard deviation, so that we have a comparison of the planned curriculum to the actual opportunities that are provided to students to learn about standard deviation in classroom settings.

6. LIMITATIONS

The findings of this study are limited by the number of participants and the immeasurable level of effort that participants put forth in designing the lesson plans. In addition, in-service teachers are known to not explain their thinking well in writing and have a significant amount of their lesson plans in their own minds only. Therefore, the in-service lesson plans may not reflect the teachers’ intended curriculum as well as the undergraduate students who are more accustomed to explaining their thoughts during college coursework. In addition, the interview data was not as rich as we had hoped. Participants often said that they did not recall why they included various activities or representations. Further, participants were anxious to discuss areas where their thinking was already well illustrated and hesitant to provide information on areas that were less defined or explained in lesson plans. Again, this seems to suggest that mathematics teacher participants lacked a degree of comfort with statistical concepts.

7. IMPLICATIONS

Although teachers in this study generally had productive pedagogical ideas about teaching standard deviation, a need for deepening participants’ knowledge of standard deviation was expressed via errors and omissions in the definition of standard deviation as well as uncertainty of how to motivate the lesson plan. It would seem that preservice mathematics teachers would greatly benefit from experiencing learning activities similar to the lesson plans described by Garfield et al. (2007) and other such activities that span Stages 1 through 6 of our framework (see Figure 1). Through such experiences, future teachers could gain a sense of the critical ideas and relationships that should be present and serve as focal points for lesson plans specific to standard deviation and variation more broadly. Such learning experiences could also help teachers recognize and resolve gaps in students’ prior experiences that create barriers to future learning. As part of such experiences, preservice teachers need access to summaries of research findings specific to learning, such as the hypothetical learning trajectory, outlined by Garfield et al. (2007), so that they can reflect upon their own learning in a broader manner that can later be useful in the act of teaching.

Given the results specific to in-service teachers, it appears that changes in pedagogy for a portion of in-service teachers are unlikely to originate without additional curricular support. Therefore, we believe that in-service teachers require curricular materials that embody the recommendations put forth by research with attention to potential learning progressions and the key connections between core statistical concepts. Many quality curricular materials are becoming available, for example *Preparing to Teach*

Mathematics with Technology (Lee, Hollebrands, & Wilson, 2010) and the Statistics Education Website (STEW) (American Statistical Association, 2010), yet they are not organized in a unit structure that show how ideas might progress over a period of time with the exception of Project-SET (2010). Without such materials and guidance, the most well-intentioned teacher may struggle to teach statistics in a manner that aligns with the Pre-K-12 *GAISE* framework and other current recommendations, such as the hypothetical learning trajectory by Garfield et al. (2007). Although encouraging results are presented here, much work and effort remains to widely equip secondary teachers with the knowledge and experiences they need to be successful teachers of statistics. In addition, careful attention must be given to what secondary students actually experience in the coming years and the learning opportunities that they are provided as many countries throughout the world attempt to implement statistics teaching in their Pre-K-12 mathematics and statistics curricula.

ACKNOWLEDGEMENTS

We would like to acknowledge the support and work of Ann Samuelson, student researcher at Drake University. We also thank the reviewers and the Associate Editor for thoughtful and insightful edits and comments.

REFERENCES

American Statistical Association. (2010). *Statistics Education Website*. [Online: http://www.amstat.org/ASA/Education/STEW/home.aspx]

Bakker, A., & Gravemeijer, K. P. E. (2004). Learning to reason about distribution. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 147–167). Dordrecht, the Netherlands: Kluwer Academic Publishers.

Common Core State Standard Initiative (CCSSI) (2010). *Common core state standards for mathematics*. Washington, DC: National Governors Association Center for Best Practices, Council of Chief State School Officers.
[Online: http://www.corestandards.org/Math/]

Confrey, J., & Makar, K. (2002). Developing secondary teachers' statistical inquiry through immersion in high-stakes accountability data. In D. Mewborn, P. Sztajn, & D. White (Eds.), *Proceedings of the 24th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education,* (Vol. 3, pp. 1267–1279). Retrieved from ERIC database. (ED471747).

delMas, R., & Liu, Y. (2005). Exploring students' conceptions of the standard deviation. *Statistics Education Research Journal, 4*(1), 55–82.
[Online: https://iase-web.org/documents/SERJ/SERJ4(1)_delMas_Liu.pdf]

delMas, R., & Liu, Y. (2007). Students' conceptual understanding of the standard deviation. In M. C. Lovett & P. Shah (Eds.). *Thinking with data* (pp. 87–116). New York: Erlbaum.

Department of Education. (2002). *Revised national curriculum statement grades R-9, mathematics.* Pretoria, South Africa: Author.

Franklin, C., Bargagliotti, A., Case, C., Kader, G., Scheaffer, R., & Spangler, D. (2015). *The statistical education of teachers.* Arlington, VA: American Statistical Association.
[Online: http://www.amstat.org/asa/files/pdfs/EDU-SET.pdf]

Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2007). *Guidelines for assessment and instruction in statistics education (GAISE) report: A Pre-K-12 curriculum framework.* Alexandria, VA: American Statistical Association.
[Online: http://www.amstat.org/asa/files/pdfs/GAISE/GAISEPreK-12_Full.pdf]

Garfield, J., & Ben-Zvi, D. (2004). Statistical literacy, reasoning and thinking: Goals, definitions, and challenges. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 3–15). Dordrecht, the Netherlands: Kluwer Academic Publishers.

Garfield, J., & Ben-Zvi, D. (2007). How students learn statistics revisited: A current review of research on teaching and learning statistics. *International Statistical Review*, *75*(3), 372–396.

Garfield, J., & Ben-Zvi, D. (2008). *Developing students' statistical reasoning: Connecting research and teaching practice*. Dordrecht, The Netherlands: Springer.

Garfield, J., delMas, R., & Chance, B. (2007). Using students' informal notions of variability to develop an understanding of formal measures of variability. In M. C. Lovett and P. Shah (Eds.), *Thinking with Data* (pp. 117–147). New York: Erlbaum.

Groth, R. E. (2013). *Teaching mathematics in grades 6–12: Developing research-based instructional practices*. Los Angeles: Sage.

Hammerman, J. K., & Rubin, A. (2004). Strategies for managing statistical complexity with new software tools. *Statistics Education Research Journal, 3*(2), 17–41.
[Online: www.stat.auckland.ac.nz/~iase/serj/SERJ3(2)_Hammerman_Rubin.pdf]

Kelly, D. G. (1994). *Introduction to probability*. New York: Macmillan Publishing.

Lee, H. S., Hollebrands, K. F., & Wilson, P. H. (2010). *Preparing to teach mathematics with technology: An integrated approach to data analysis and probability*. Dubuque, IA: Kendall Hunt.

Lehrer, R., Kim, M., & Schauble, L. (2007). Supporting the development of conceptions of statistics by engaging students in measuring and modeling variability. *International Journal of Computer form Mathematical Learning, 12*, 195–216.

Makar, K., & Confrey, J. (2005). "Variation-talk": Articulating meaning in statistics. *Statistics Education Research Journal, 4*(1), 27–54.
[Online: https://iase-web.org/documents/SERJ/SERJ4(1)_Makar_Confrey.pdf]

McClain, K., McGatha, M., & Hodge, L. (2002). Improving data analysis through discourse. *Mathematics Teaching in the Middle School, 5*(8), 548–553.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. London: Sage.

Ministério da Educação (2006). *National curricular guidelines for secondary school: Nature sciences, mathematics, and their technologies*. Brasilia, Brazil: Author.

Mlodinow, L. (2008). *The Drunkard's walk: How randomness rules our lives*. New York: Pantheon Books.
[Online: http://www.corestandards.org/Math/]

Peters, S. (2014). Developing understanding of statistical variation: Secondary statistics teachers' perceptions and recollections of learning factors. *Journal of Mathematics Teacher Education, 17*(6), 539–582.

Project-SET (2010). *Project-SET: Statistics Education for Teachers*.
[Online: https://project-set.com]

Reading, C. (2004). Student description of variation while working with weather data. *Statistics Education Research Journal*, *3*(2), 84–105.
[Online: https://iase-web.org/documents/SERJ/SERJ3(2)_Reading.pdf]

Reading, C., & Shaughnessy, M. (2004). Reason about variation. In D. Ben-Zvi, & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 201–226). Dordrecht, the Netherlands: Kluwer Academic Publishers.

Rossman, A. J., & Chance, B. C. (2011). *Workshop statistics: Discovery with data* (4th ed.). Hoboken, NJ: John Wiley & Sons.

Sánchez, E., Silva, C. B., & Coutinho, C. (2011). Teachers' understanding of variation. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics: Challenges for teaching and teacher education. A Joint ICMI/IASE study: The 18th ICMI study* (pp. 211–221). New York: Springer.

Senior Secondary Board of South Australia. (2002). *Mathematical studies curriculum statement.* Adelaide, Australia: Author.

Shaughnessy, J. M. (2007). Research on statistics learning and reasoning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 957–1010). Reston, VA: National Council of Teachers of Mathematics.

Silva, C. B., & Coutinho, C. (2008). *Reasoning about variation of a univariate distribution: A study with secondary mathematics teachers.* In C. Batanero, G. Burrill, C. Reading, & A. Rossman (Eds.) *Joint ICMI/IASE Study: Teaching statistics in school mathematic. Challenges for teaching and teacher education. Proceedings of the ICMI Study 18 and 2008 IASE Roundtable Conference.*
[Online: https://iase-web.org/documents/papers/rt2008/T2P7_Silva.pdf]

Smith, M. S., & Stein, M. K. (2011). *5 practices for orchestrating productive mathematics discussions.* Reston, VA: National Council of Teachers of Mathematics.

Stein, M. K., Remillard, J., & Smith, M. S. (2007). How curriculum influences student learning. In F. K. Lester, Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 319–369). Reston, VA: National Council of Teachers of Mathematics.

Sullivan, M., III. (2010). *Statistics: Informed decisions using data* (3rd ed.). Wright City, MO: Pearson.

Tarr, J. E., Chávez, O., Reys, R. E., & Reys, B. J. (2006). From the written to enacted curricula: The intermediary role of middle school mathematics teachers in shaping students' opportunities to learn. *School Science and Mathematics, 106*(4), 191–201.

Tarr, J. E., & Jones, D. L. (2010). Recommendations for statistics and probability in school mathematics over the past century. In B. J. Reys, R. E. Reys, & R. Rubenstein (Eds.), *Mathematics curriculum: Issues, trends, and future direction: Seventy-second Yearbook* (pp. 65–76). Reston, VA: National Council of Teachers of Mathematics.

The Common Core Standards Writing Team. (2012). *Progression for the Common Core: High school statistics and probability.*
[Online: commoncoretools.me/wp-content/uploads/2012/06/ccss_progression_sp_hs_2012_04_21_bis.pdf]

Watson, J. M., & Moritz, J. B. (2000). Developing concepts of sampling. *Journal for Research in Mathematics Education*, *31*, 44–70.

Watson, J. M., Kelly, B., Callingham, R., & Shaughnessy, M. (2003). The measurement of school students; understanding of statistical variation. *International Journal of Mathematical Education in Science and Technology, 34*(1), 1–29.

MARYANN E. HUEY
Drake University
Department of Mathematics
2507 University Ave.
Des Moines, IA 50322
USA

APPENDIX: LESSON PLAN ASSIGNMENT PROMPT

Write a lesson plan for which the objective is to have students understand standard deviation. The lesson plan can be contained within a regular class, a block period, or run over a few days.

Your primary goal is to describe in detail the key ideas that you would develop during this period, and the key tasks/activities/assignments that you would include to make sure students develop a solid understanding about standard deviation. Focus on describing the essential aspects of the lesson (i.e., don't detail going over homework for the first 5 minutes – unless it's an essential part of developing the key ideas you have in mind).

You can assume that the students in your class have already learned about the mean, interquartile range (IQR), and mean absolute deviation (MAD) for data distributions. Students are transitioning from grade 6 measures of variability in the CCSS-M (MAD and IQR) to standard deviation. Below are a few of the high school standards about standard deviation:

- CCSS.Math.Content.HSS-ID.A.2. Use statistics appropriate to the shape of the data distribution to compare center (median, mean) and spread (interquartile range, standard deviation) of two or more different data sets.
- CCSS.Math.Content.HSS-ID.A.3. Interpret differences in shape, center, and spread in the context of the data sets, accounting for possible effects of extreme data points (outliers).
- CCSS.Math.Content.HSS-ID.A.4. Use the mean and standard deviation of a data set to fit it to a normal distribution and to estimate population percentages. Recognize that there are data sets for which such a procedure is not appropriate. Use calculators, spreadsheets, and tables to estimate areas under the normal curve.