The Big Ideas in the statistics education of our students: Which ones are the biggest?

J. Michael **Shaughnessy**
Portland State University
United States of America
mikesh@pdx.edu

Summary

Five decades of research and curriculum development on the teaching and learning of statistics have produced many recommendations from both researchers and national organizations for the statistical education of our students as well as how statistics should be taught. Within the last ten years work by both statisticians and statistics educators has been focusing more on a collection of big ideas that are the most important concepts and processes to develop statistical thinking for our students, for our work force, and for the lifelong statistical literacy of our citizens. In this paper I look back at the roots of big ideas in statistics education and then identify what I believe are the two most important overarching ideas that can orchestrate the statistical education of our students starting from the elementary years into tertiary. The paper includes research on student thinking about big ideas in statistics and recommendations for the future of teaching and research in statistics education.

Key Words: Statistics education, distribution, inference, variability, expectation, sampling, statistical investigation processes.

Recommendations for the Big Ideas in Statistics Education: A Retrospective

Prior to the 1960's there was almost no statistics included in the school curricula of the nations of the world. In their historical review What is Statistics Education? Zeiffler, Garfield, and Fry (2018) point to recommendations starting in the 1960's in which several curriculum projects in the UK recommended the inclusion of probability and statistics in schools for students ages 11 – 16. In 1967 the American Statistical Association (ASA) and the National Council of Teachers of Mathematics created the Joint Committee on the Curriculum in Statistics and Probability in the U.S. and Canada. The Joint Committee began to spearhead the publishing of materials for teaching statistics in the early 1970's such as *Statistics: A Guide to the Unknown* (Tanur, Mosteller, Kruskal, Link, Pieters, & Rising, 1972), and *Statistics by Example* (Mosteller, Kruskal, Link, Pieters, & Rising, 1973). The Joint Committee has continued to sponsor and promote statistics education and the professional development of teachers to this day with curriculum materials such as *The Quantitative Literacy Project* (Ganadesikan et. al., 1995) and the recommendations for the teaching and learning of statistics in the GAISE documents, *Guidelines for Assessment and Instruction in Statistics Education* (Franklin, Kader, Mewborne, Moreno, Peck, Perry, & Schaeffer, 2007).

Early attempts to include statistics in the education of school age students prompted research into the teaching and learning of statistics which began in the1970's (particularly in the UK, Germany, Israel, and the U.S. For details see Shaughnessy, (1992)). The growing international interest in teaching and research in statistics education eventually gave birth to the *First International Conference on Teaching Statistics* in Sheffield, England in 1982, ICOTS I. Since that time an ICOTS has been convened every four years up to the most recent ICOTS X which was held in Hiroshima in 2018. When the next ICOTS is held in Rosario, Argentina in 2022, an ICOTS will have been held on every continent, and in 11 different countries.

NCTM Standards for Statistics Education of K-12 Students

Over the fifty years since the birth of statistics education as a discipline both the research and practitioner communities have been continually honing in on the most important ideas in statistics for our students and citizens to know and be aware of. Starting with its *Agenda for Action* document (NCTM, 1980), and subsequently with its *Curriculum and Evaluation Standards for School Mathematics* (NCTM, 1989)*,* the National Council of Teachers of Mathematics began to advocate for specific recommendations for teaching statistics to Grades K–12 in the United States and Canada. The 1989 standards were NCTM's first foray into establishing goals for school mathematics. With regard to statistics, NCTM included the following recommendations:

For grades K – 4:

- Formulate and solve problems that involve collecting, describing and analyzing data
- Construct, read and interpret displays of data
- Explore the concepts of chance

For grades 5—8:

- Systematically collect, organize and describe data
- Construct, read and interpret tables, charts, and graphs
- Make inferences and convincing arguments, and evaluate the arguments of others based on data analysis
- Develop an appreciation for statistical methods as powerful means for decision making

For grades 9 – 12

- Construct and draw inferences from charts, tables and graphs that summarize data from real world situations
- Use curve fitting to predict data
- Understand and apply measures of central tendency, variability, and correlation
- Understand sampling and recognize its role in statistical claims
- Design a statistical experiment to study a problem
- Analyze the effects of data transformations on measure of center and variability
- Test hypotheses using appropriate statistics

The 1989 standards started with a data analysis perspective in grades K–8, but took quite a jump in depth and abstraction in grades 9–12 with statistical design, mathematical transformations on parameters, and traditional hypothesis testing. Many introductory college instructors would probably be quite happy if their students mastered these original NCTM grade 9–12 statistics standards. These 1989 standards are predominantly a list of content, concepts and procedures that students should know and be able to do. However, the process of making inferences is included for grades 5–12. It will be interesting to look back at these 1989 NCTM standards from the point of view of later recommendations about the big ideas in statistics.

Ten years later NCTM produced its second round of standards recommendations in *Principles and Standards for School Mathematics* (*PSSM*), (NCTM, 2000). This time NCTM's standards for statistics were organized under four broad big ideas at grades K–2, 3–5, 6–8, and 9–12, with additional lists that further explicate each of these four expectations at each grade level. *PSSM* recommended these that instructional programs from prekindergarten through grade 12 should enable all students to do the following in Data Analysis and Probability:

- Formulate questions that can be addressed with data and collect, organize and display relevant data to answer them
- Select and use appropriate statistical methods to analyze data
- Develop and evaluate inferences and predictions based on data
- Understand and apply basic concepts of probability

This time NCTM organized the big ideas in statistics education more from a statistical processes point of view than just a list of concepts and statistics content as in 1989. While the organizational headings remain the same throughout all for grade bands, the sophistication in addressing the big ideas of course grows up through the later grades. For example, NCTM presents this trajectory across the grade bands for the process of developing and evaluating inferences and predictions based on data:

PreK–2. Discuss events related to students' experiences as likely or unlikely

Grades 3–5. Propose and justify conclusions and predictions based on data and design studies to further investigate conclusions or predictions

Grades 6–8. Use observations about differences between two of more samples to make conjectures about the populations from which the samples were taken.
Make conjectures about possible relationships between two characteristics of a sample on the basis of scatterplots and approximate lines of fit.
Use conjectures to formulate new qu3estions and plane new studies to answer them.

Grades 9–12. Use simulations to explore the variability of sample statistics from a known population and to construct sampling distributions.
Understand how sample statistics reflect the values of population parameters and use sampling distributions as the basis of informal inference.
Evaluate published reports that are based on data by examining the design of the study, the appropriateness of the data analysis, and the validity of the conclusions.
Understand how basic statistical techniques are used to monitor process characteristics in the workplace.

Notable in the *PSSM* standards when compared to the earlier standards is the growing emphasis and detail on making and testing data-based conjectures and the introduction of the term "distribution", especially in regard to creating sampling distributions via simulations and using them to make inferences about populations.

The Central Role of Variability—Recommendations from the American Statistical Association

During the 1980's and 1990's much of the statistics education in school curricula concentrated on measures of center and mostly neglected the important role that variability plays in statistics. This was particularly true in elementary school mathematics curricula that introduced statistics to students by calculating mode, median, and mean. In his position paper on statistics content and pedagogy, David Moore (1997) a president of the ASA emphasized the crucial role that variability plays in statistics education, without variability statistics would not even exist. The writings of Moore and others sounded a clarion call for rethinking what the big ideas in statistics education really were. Subsequently statistics education researchers began to concentrate more on investigating students reasoning about variability. (See for example Shaughnessy, Watson, Moritz & Reading, 1999; Melitou, 2002; Toruk & Watson, 2000; Watson, Kelly, Callingham & Shaughnessy, 2003; Watson & Kelly, 2004; Reading & Shaughnessy, 2004). Thus, when the GAISE document (*Guidelines for Assessment and Instruction in Statistics Education)* was published the American Statistical Association (Franklin et al, 2007), it appropriately highlighted variability as a central organizing idea of statistics education. The GAISE recommendations focus on teaching statistics through the cycle of four components of the statistical investigation process while highlighting the important role that variability plays in each step of the process.

I. Identify a Statistical Question—*Anticipate Variability,*
formulate questions that can be answered with data

II. Collect Data—*Acknowledge Variability,*
Design for differences and plan to collect appropriate data

III. Analyze Data—*Account for Variability by Using Distributions*
--select appropriate graphical and numerical methods to analyze the data

IV. Interpret Results—*Allow for Variability as you look beyond the data*
--interpret the analysis and relate it back to the original question

Figure 1. The statistical investigation process.

The GAISE for statistics education chose to focus on statistical processes rather than a list of specific statistical concepts and procedures as the organizing principles for teaching statistics in K–12. GAISE describes three levels (A, B, and C) of sophistication and growth for each of these four components in the statistical investigation cycle. The three levels roughly correspond to recommendations for grades K–4, 5–8 and 9–12 respectively.

Recommendations for the big ideas from research on the teaching and learning of statistics

Along with the increased attention to statistics education in schools around the world and the accompanying recommendations for pursuing the big ideas in statistics the research literature in statistics education has grown exponentially from its initial roots in the 1970's. There are opportunities to present research in statistics education at international conference such as the ICOTS conferences and *Psychology and Mathematics Education* (*PME),* national conferences such as the *US Conference on Teaching Statistics (USCOTS),* the NCTM *Research Conference,* the *Mathematics Education Research Group of Australasia (MERGA),* and *Conferencia*

Interamericana Educacíon Matematica CIAEM, as well as many other mathematics education conferences throughout the world. Statistics education research has several dedicated journals for publications including the *Statistics Education Research Journal (SERJ)* and the *Journal of Statistics Education* (JSE). *SERJ* is published by the *International Association of Statistics Education (IASE)*, a branch of the *International Statistics Institute (ISI), JSE* is published by the ASA. In addition, journals that are predominantly in mathematics education such as *The Journal of Research in Mathematics Education (JRME), Mathematical Thinking and Learning (MTL),* and *Educational Studies in Mathematics (ESM)* as well as many other international journals have routinely included articles on research in statistics education. Research in statistics education has been reviewed and synthesized over the years and numerous recommendations for the teaching and learning of statistics have been cited in those reviews (Shaughnessy; 1992, Shaughnessy, Garfield, and Greer, 1996; Ben-Zvi and Garfield, 2004; Shaughnessy, 2007; Garfield & Ben Zvi, 2008; Langrall, Makar, Nilsson, & Shaughnessy, 2017; Biehler, Frischmeier, Reading & Shaughnessy, 2018). Among the Big Ideas in statistics for recommended for practitioners to concentrate on are those from Garfield & Ben-Zvi (2008) and Watson, Fitzallen & Carter (2013). In their book *Developing Students' Statistical Reasoning* Garfield and Ben-Zvi argue that the teaching of statistics at all levels should concentrate on nine central concepts:

- Data
- Distribution
- Variability
- Center
- Statistical Models
- Randomness
- Co-variation
- Sampling
- Inference

Many of the groups that made recommendations about the big ideas in statistics prior to Garfield & Ben- Zvi concentrated on important statistical processes such as posing statistical questions, gathering data, and analyzing data. Garfield & Ben-Zvi returned to an emphasis on the most important statistical concepts for students and teachers in statistics education to address. It is clear that any lists for the big ideas in statistics education must consider both statistical processes and statistical concepts. These two perspectives on organizing what is most important in statistics interact with and support one another. Watson et al (2013) identified five big ideas in statistics that they modeled as the vertices of a pentagon with edges connecting each pair of vertices (resulting in a pentagram). Their five Big Ideas for statistics education are variation, expectation, distribution, randomness, and inference. To the notion of a distribution Watson et al (2013) added the importance of making inferences from distributions of data and from comparisons across several distributions of data. They thus tied distributions to the realm of hypothesis testing, starting at the level of informal inference, and anchored in expectation and variation in the data. Is a distribution of data predominantly due to chance alone, or is something else accounting for the aspects of the distribution such as the shape, the variability, or clustering of the data? It is interesting to compare and contrast Watson et al.'s approach to big ideas with that of Garfield & Ben-Zvi. There is agreement between these sets of authors that variation, distribution, randomness and inference are among the big ideas in statistics. Watson et al use the term expectation, while Garfield and Ben-Zvi prefer to refer to use center. Although the word center may evoke notions of procedures like calculating means and medians, I suspect that Garfield & Ben-Zvi were thinking of center in a wider sense, similar to the term expectation used by Watson. Watson's model of the big ideas in statistics is informed by her research on the tension that students experience between variation and expectation when they make predictions or analyze data sets (Watson, 2009).

NCTM's Essential Understandings recommendations for teachers of statistics

As the implementation of more statistics has grown in school mathematics programs, many of our mathematics teachers have found themselves in the position of having to teach statistics concepts when they have little or no preparation in statistics themselves. In order to provide some professional development and assist middle and secondary school mathematics teachers in adding statistics to their teaching repertoire the National Council of Teachers of Mathematics included statistics in their series on the *Essential Understandings* in school mathematics. (The *Essential Understandings* books cover algebra, geometry, number and operation, proportional reasoning, and mathematical reasoning as well as statistics.) Both *Essential Understanding of Statistics Grades 6–8* (Kader & Jacobbe, 2013) and *Essential Understanding of Statistics Grade 9–12* (Peck, Gould, & Miller, 2013) identify the big ideas in statistics that all teachers should know and be able to teach at their respective grade levels. Both books include copious sample tasks for teachers to explore the big ideas themselves, and to use in teaching their students. In the grade 6–8 book, Kader & Jacobbe identify four big ideas for teaching statistics to middle school students:

- Variability in Data and Distributions
- Comparing Distributions
- Associations between Two Variables
- Samples and Populations

The concept of a distribution plays a prominent role in all four of these recommendations if one considers that bivariate distributions of data form the basis for exploring associations between variables. Peck et al describe a collection of statistical processes that form the foundation of the big ideas in the grade 9–12 *Essential Understandings of Statistics* book. They are especially interested in strengthening our teachers' knowledge and comfort with these ideas:

1. Data consists of structure and variability
2. Distributions describe variability
3. Hypothesis tests answer the question, "Do I think this could have happened by chance?"
4. The way data are collected matters.
5. Evaluating an estimator involves considering bias, precision, and sampling method.

For Peck et al these five organizational big ideas are interrelated. Statistical models describe and account for variability in both populations and in the values of sample statistics depicted in sampling distributions. Hypothesis testing is the basis for making decisions under uncertainty based on the limitations of the data provided. The data upon which statistical decisions are made are only as good as the care with which they are produced, so that attention to sources of bias and precision in estimating parameters such as measures of center and variation is critical. Peck et al describe finer grain detail for the essential understandings within each of these five big ideas and provide examples for teachers to consider for their own understanding, as well as for use with their students.

The two BIGGEST ideas in statistics education

Suppose that you were asked to pick two ideas in statistics that you thought were the most important ones for our students to learn and our citizens to be competent in understanding. What would be your choice? What are your two BIGGEST ideas in statistics education? The most important goal for statistics education is to enable our students and citizens to understand that making decisions under conditions of uncertainty is based upon samples of data. We rarely have access to information about the entire population under consideration when making statistical decisions or estimating the likelihood that some event occurred by chance alone. Statistics does not appeal to mathematical proof based on deterministic reasoning anchored in axiomatics. Statistics involves making decisions that we perceive are most likely to be true based on data that is generated under conditions of uncertainty. Given that the pre-eminent goal of statistics education is to understand decision making under uncertainty, I claim that the two biggest ideas in statistics education are *distribution* and *inference* because these two ideas are the heart and soul of statistical decision making. I base this conclusion in part on the analysis above of the trajectory and development of the recommendations of various organizations and groups of researchers throughout the history of statistics education, but also on some more recent research in statistics education that gives added support to the claim that *distribution* and *inference* are the two biggest overarching ideas in statistics education.

Examples of research on students' reasoning about big ideas in statistics

Beginning in the 1990's researchers began to investigate students' understandings of big statistics concepts from a developmental perspective. Research on student understanding of concepts such as average and variability has found trajectories of student reasoning, levels of student understanding that become deeper over time. Furthermore, concepts such as expectation and variation are components of bigger ideas such as distribution and inference. Examples of some of these reasoning trajectories of big ideas in statistics from research are discussed below.

Expectation.

The term expectation encompasses research on measures of center such as mean, median, and mode as well as considerations of any expected clumping in the data, as many distributions of data have bimodal or even multimodal characteristics. Mokros & Russell (1995) discovered one of the first trajectories of students' understanding of average. Using interview tasks that involved "messy situations" from everyday familiar contexts with grades 4, 6, and 8 students Mokros & Russell identified five different ways that students thought about average: average as *mode* (mosts), average as *algorithm*, average *as reasonable*, average as *midpoint*, and average as *balance point*. Watson & Moritz (2000) interviewed about a hundred students grades 3, 5, & 7 over time and found their conceptions of average moved from telling idiosyncratic stories about average to thinking of average as 'mosts or middles', and eventually to average as representative of a data set. Reflecting upon the research on students' conceptions about expectation, Konold & Pollatsek (2002) proposed that students' thinking about the mean includes average as *typical value,* average as *fair share,* average as *data reducer,* and average as *signal amid noise.* It is clear from the research on students reasoning about expectation that students possess a rich collection of conceptions about centers which teachers can build upon. (For a more detailed discussion about research on students' conceptions of average, see for example Shaughnessy, 2007).

Variation.

Three developmental frameworks for students reasoning about variability are presented and compared by Langrall et al (2018, p. 494) in the NCTM *Compendium of Research in*

Mathematics Education. In his framework Ben-Zvi (2004) notices at first students just recognize variability across various data values. Later students use variability to compare groups, next they combine measures of spread and center to compare groups, and eventually they consider variability as a construct within and between distributions of data. Watson, Callingham, & Kelly (2007) describe a progression of student thinking that encompasses both expectation and variation. Students first acknowledge just one or the other, then recognize that both expectation and variation are present in data, later on students begin to see an interaction between centers and variability as they develop proportional reasoning. Reid & Reading (2008) investigate students' consideration of variation in data and describe a hierarchy of student reasoning ranging from no consideration of variation, to recognition of variation within a group, to recognition of variation between groups that leads to considering inference. In an analysis of the research on students' conceptions of variability, Shaughnessy (2007) outlined eight types of conceptions of variability that were identified in research including:

1. Variability in *particular values* in a data set
2. Variability as *change over time*
3. Variability as the *whole range* of a data set
4. Variability as the *likely range* of a sample
5. Variability as *distance from some fixed point*
6. Variability as *sum of residuals*
7. Variability as *covariation or association*
8. Variability as *distribution*

The first four of types of variability in this list involve an exploratory data analysis perspective, while the last four types refer primarily to ways to measure variability. The terms variability and variation are sometimes used almost interchangeably, however some authors (e.g., Reading and Shaughnessy, 2004) prefer to use the term variability as the tendency for a characteristic to change while the word variation refers to a measurement of a changing characteristic. Thus, the first four types above refer to variability, while the last four involve variation, some type of measurement of change. Research on type 4, variability as the *likely range of a sample*, led to further research on students' conceptions of sampling distributions. A closer look at some research tasks and student responses to those tasks may provide insight into why distribution is one of the two biggest ideas in statistics education.

100 candies, 20 yellow, 50 red, and 30 blue, are put in a jar and mixed together. A student pulls 10 candies from the mixture, counts the number of reds, and writes that number on the board. Then the student puts the candies back in the bowl and mixes them all up again. Four more students also draw a sample of 10 candies, and write their number of reds on the board. What numbers would you predict for the number of reds in each of those five samples of 10 candies? Write your predictions in the spaces below.

_________ _________ _________ _________ _________

Why do you think those would be the numbers of reds in the five samples?
Fig 2. The candy sampling task.

This task and variations of it were given to hundreds of students in grades 4, 5, 6, 9, & 12 in Australia, New Zealand, and the United States (Shaughnessy, Watson, Moritz, & Reading, 1999). The task was used to determine what students perceived as the *likely range* of values that would occur in a repeated sampling scenario. Student responses fell into clusters that were deemed *narrow, wide, high, low,* and *reasonable.* For example, some students said they'd expect the results to be 6, 7, 5, 8, 9 because 'there are a lot of red in the jar.' This is typical of a *high* response as all the sample predictions are above the expected value of 5 red. High responses were based on thinking about 'mosts', not about the proportion of reds in the mixture. In *low* responses like 3, 4 ,3, 5, 2 students felt that the other colors would overwhelm red, as there were two of them. Students who predicted *wide*, like 1, 5, 7, 10, 2, did so because 'anything can happen.' On the other hand, some students predicted results like 5, 5, 6, 5, 6, or even 5, 5, 5, 5, 5 because 'that's what is supposed to happen.' Such *narrow* predictions put too much weight on the theoretical probability of obtaining 5 red candies on any one pull, and neglected the possibility of variability in outcomes. Overall, some students attended to centers too much, some to variability too much, and some students did predict a *reasonable* range of sample outcomes around the center with predictions such as 3, 7, 5, 6, 4. Research on tasks such as the candy sampling task helped to spawn further research into students' conceptions about distributions, in particular their conceptions about sampling distributions. The candy sampling task also points to the tension that can arise between attending to expectation and attending to variability in data, especially when students are asked to make predictions for samples (Watson & Kelly, 2004).

Distribution.

Comparisons across several hypothesized developmental frameworks of the concept of distribution are provided in Langrall et al (2018, p. 494). In each of these developmental frameworks the researchers acknowledge that the idea of distribution in statistics encompasses the aspects of shape, variability, and expectation, and that integration of all of these aspects is required for students to be able to reason about and make inferences from distributions. Reading and Reid's framework (2006) for understanding distributions starts with students acknowledging one parameter of a distribution, then several parameters, then integrating centers and spreads when considering data aggregates, and finally to a second cycle of reasoning development which involves students' growth in making inferences from distributions. A framework for distributional thinking was proposed by Noll & Shaughnessy (2012) based on their research on students' conceptions of sampling distributions (See Figure 3). According to Noll & Shaughnessy, students' development of the concept of distribution involves the gradual integration of shape, centers, and spread. Students at first notice them as individual aspects of a distribution, then learn to make predictions for sampling distributions by relying on both expectation and variation. Students' reasoning progresses through four levels from *additive* to *proportional* and finally to *distributional* reasoning in the progression identified by Noll & Shaughnessy.

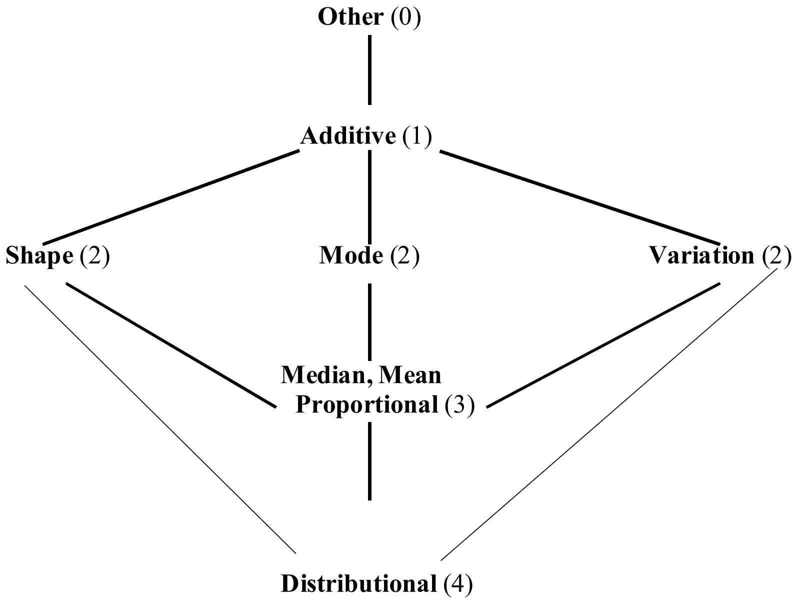


Figure 3. Lattice of student reasoning when making predictions about sampling distributions.

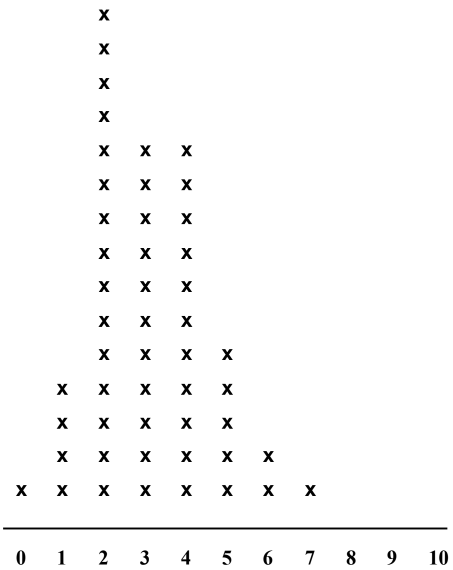
The lattice indicates the development of conceptions of expectation from 'more' to 'most' (mode) to 'means and medians' which involves proportional reasoning, and finally to reasoning about distributions which includes the coordination of aspects of shape, expectation, and variation. Student responses to tasks like the *Prediction Task* and the *Mystery Mixture* task led Noll & Shaughnessy to the development of this reasoning lattice. A version of each of these tasks is presented here.

Working in small groups, students in a class pull samples of 10 candies from a jar that has 1000 candies. They pull 50 samples of ten. The jar has 250 yellow and 750 red candies in it. Each time they put the sample back and remix the jar.
Consider the number of reds in each handful. Where would you expect 95% of the handfuls of ten candies to be?

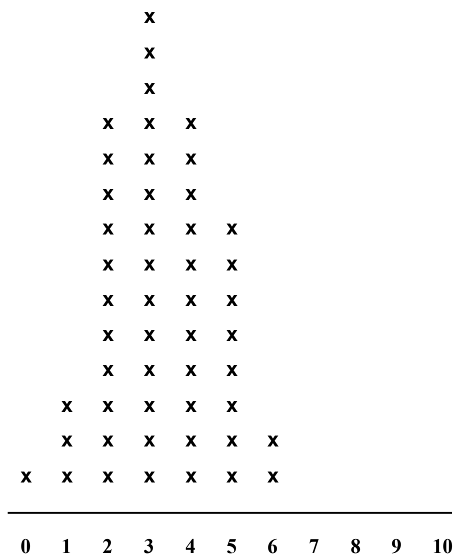
From __________ # reds To _______________# reds (Fill in the blank spaces).
Why do you think that?
Complete the frequency chart below to show what you think the numbers of reds in 50 trial handfuls might look like. (Note: students were provided labeled graph paper).

Figure 4. The Prediction Task

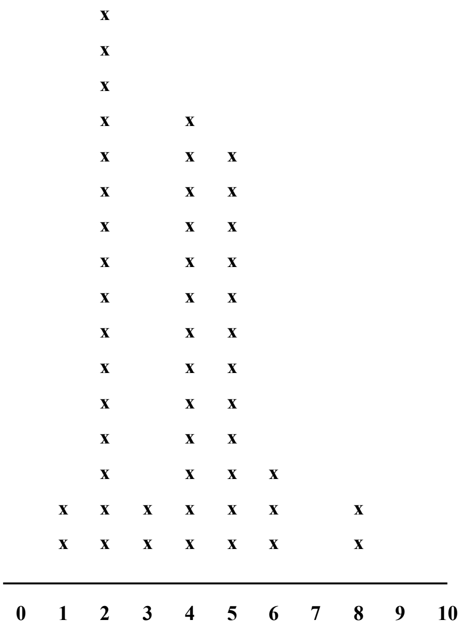
The four frequency graphs below all came from a class that is trying to estimate a mystery mixture of 1000 red and yellow candies in a large jar. They pulled 50 samples of size 10, recorded the number of reds in each sample, and then replaced and remixed each time.



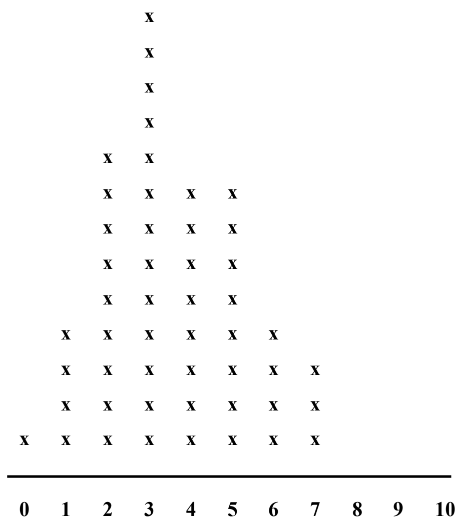
(Graph 1)



(Graph 2)



(Graph 3)



(Graph 4)

a) What do you think the mixture in the jar might be? b) Explain why you think this.

Figure 5. The Mystery Mixture Task

In the Prediction task students are given the proportions of colors in the parent population and then asked to predict what a sampling distribution for sample proportions will look like. Most students' responses fell into the prediction categories *wide, narrow, and reasonable* for both their predicted range for the # of red in handfuls, and in the graphs that they constructed for 50 sample proportions. In the Mystery mixture task students are provided with the results from multiple sampling distributions and asked to use this information to infer what the parent population is. Students who predicted 200-250 reds in the mystery mixture reasoned using a 'mosts' point of view, and were usually heavily influenced by the graphs of distributions 1 & 3 which have a mode of 2 reds in handfuls. Student who looked for 'balance points' of the graphs reasoned proportionally and tended to predict around 300 reds in the mixture. Still other students noticed that the graphs tended to be skewed to the right, they integrated shape as well as center and spread into their reasoning and inferred that the mixture likely had more than 300, perhaps even 350 - 400 reds in it. In both the Prediction task and the Mystery Mixture task students had opportunities either to focus solely on one of the aspects of the sampling distributions (shape, centers, variability) or to integrate them into their predictions and inferences.

Inference.

The Mystery Mixture task above is an example in which students are asked to make an inference beyond the data at hand, in this case beyond the given sampling distributions of a sample statistic. There is no formal hypothesis testing here, students are simply asked what they believe the composition of the parent population is and why they believe it. Researchers and curriculum developers now refer to this type of inference as 'informal inference'. Informal inference has its roots in exploratory data analysis, often in the exploration of data that have been produced from simulations. Students can estimate likelihoods from samples of data without resorting to a test statistic or working with a probability distribution. Cobb (2007) argued that introductory statistics courses should start with inference early on prior to any hypothesis testing that resorts to theoretical distributions. Since the logic of formal statistical inference has always been a difficult stumbling block for statistics students, educators have been experimenting with various approaches to the early introduction of inference that avoid some of the cognitive complexity and pitfalls in formal inference. Rossman (2008) promotes introducing simulations of randomization tests as a more transparent informal approach to statistical inference. Zeiffler, Garfield, Del Mas, & Reading, (2008) define informal statistical inference as students using their informal statistical knowledge about observed samples to make arguments to support inferences about unknown populations. Makar and Rubin (2018) point out that there is general agreement that the important characteristics of informal inference are: i) a claim is made that goes beyond the data at hand; ii) the data are used as evidence to support the claim; iii) the claim involves the articulation of uncertainty—estimated likelihoods or probabilities are involved; iv) decisions or inferences are based upon aggregates in the data, variability, or shape so decisions are based upon aspects of distributions of data; v) contextual knowledge plays a role in the analysis and inference integrated with the information in the sample of data. Makar and Rubin share examples from research with both elementary and middle grade students who reasoned about tasks and made informal inferences from data that they had collected. Inference is one of the two biggest ideas in statistics because students even at a very young age can begin to make inferences based on data that they have collected on a statistical question that they themselves posed. The statistical investigation cycle outlined in the GAISE document—pose a question, collect data, analyze the data, make conclusions—is even more powerful for students when it includes making inferences in the analysis and conclusions phases.

Recommendations for future research and teaching the big ideas in statistics

The future of research in statistics education

Many of the big conceptual ideas in statistics such as distribution, expectation, variation, and randomness identified above reside predominantly in the third stage of the GAISE statistical investigation cycle, in the stage of analyzing the data. However, in addition to these conceptual big ideas, there are process big ideas in the statistical investigation cycle, such as posing a statistical question, generating appropriate data to answer a statistical question and communicating the results and conclusions of the investigation to others. These big statistical processes ideas must also be included in the statistical education of our students, as well as in the professional development of our classroom teachers. Unlike the big conceptual ideas such as expectation, variability, and distribution there has not been much research into developing students' ability to pose statistical questions or on their ability to communicate and defend their their inferential conclusions. The first and last phases of the statistical investigation cycle have not yet been adequately explored by research, especially when compared to how much research has been conducted on the the collecting and analyzing phases. The ASA conducts a Statistics Poster contest each year at the elementary, middle, and secondary student levels. The poster contest could provide a fertile ground for research on what students learn from their involvement in the statistical investigation cycle. Posters provide information on both the statistical question posed, and the results communicated. More research is needed on students' thinking processes as they pose a statistical question,and as they communicate their results. Overall, more research into student thinking about both the concepts and processes of the entire statistical investigation cycle would be beneficial to both the teaching and research communities.

Over the last twenty-five years research has concentrated on particular areas and obtained results robust enough to support the existence of developmental frameworks for students reasoning about expectation, variation, and distribution (Langrall et al, 2017). More research is needed to further validate the developmental frameworks that have already been proposed. Meanwhile the next 'big idea' in this research progression appears to be inference, in particular informal inference. A special issue of the *Statistics Education Research Journal* was dedicated to articles about informal inference, particularly within statistical modeling contexts (*SERJ* 16 (3), November, 2017). Various definitions of informal inference have been proposed and some of the components of informal inference have been identified. However, a developmental framework for students' reasoning about inference analogous to those for variability and distribution does not yet exist. Case & Jacobbe (2018) report a framework for understanding students' difficulties when making inferences from simulations. However, much more research is needed about how inferential reasoning develops starting with young children and up through the grades in order to identify potential levels of student reasoning about inference.

The future of teaching statistics

Teaching is a social process, it involves countless interactions between students and their instructors. Any recommendations for teaching statistics must include considerations about the teacher as well as the students. Our teachers are on the front line of statistics education, and many of the teachers in our current work force do not have very much experience with statistics, much less actually teaching statistics. Most of them are mathematics teachers. As Cobb & Moore (1997) pointed out so well, mathematics and statistics are very different disciplines. Mathematics is grounded in certainty, deductive reasoning based on assumed axioms and previously established results builds toward new certain truths. Statistics on the other hand lives in the realm of uncertainty, statistical results are couched in terms of likelihoods, probabilities, confidence intervals. Mathematics and statistics are epistemologically and philosophically different from one another. Our teachers need experiences themselves in carrying out statistical investigations—perhaps in conjunction with their students—in order to immerse themselves in thinking about drawing conclusions from data. In this regard, teachers need to develop both their content knowledge and their pedagogical content knowledge of statistics. The NCTM *Essential Understandings* series includes books on the big ideas in statistics that can provide some content knowledge support for middle and secondary school teachers of statistics (Kader & Jacobbe, 2013; Peck et al, 2013). The ASA GAISE document provides support for developing teachers' pedagogical content knowledge in statistics. It provides many examples of tasks that can be implemented in the classroom using the statistical investigation cycle, and outlines a progression of levels of student understanding about the statistical investigation cycle.

As for our future teachers of statistics, the ASA recently developed and published *The Statistical Education of Teachers (SET)* which lays out recommendations for the statistical education of all perspective teachers, elementary, middle, and secondary teachers alike (Franklin, Bargagliotti, Case, Kader, Scheaffer, & Spangler, 2015). *SET* recommends both coursework and statistical modeling experiences for all teachers. Coursework should begin with a first course that using a data analytic approach (all teachers), and the recommendations for middle and secondary school teachers include additional coursework in statistical methods and statistical modeling. The ASA has taken a very futuristic view in the *SET* document that projects that the need for statistics education will continue to grow throughout the world, and that our teaching profession will need to know much more about statistics and statistical thinking in the future to prepare students and citizens to better understand and be able to work in our data intense world.

What about our students, our learners, our future workers and citizens? What does our walk through the history and research on Big Ideas in statistics suggest for the teaching and learning of statistics?

Start with the big ideas. Right away introduce the concept of a statistical question, a question that requires data to be answered. Make sure that students understand the difference between mathematics and statistics, that they are two different disciplines, that they involve two different types of reasoning. Give students opportunities to reason about distributions of data and to make informal inferences early on. Provide students with multivariate data sets to explore and have them collect multivariate data themselves and then ask them, "What do you notice? What do you wonder about?" Get students to make comparisons between data sets, across distributions, early on. Get students involved in generating sampling distributions from repeated samples from both known and unknown populations, and then making informal inferences from the samples they've obtained. Have students recognize and attend to the important aspects of distributions such as shape, centers, and variability, then begin to introduce ways of measuring expectation and variation. Use the developmental frameworks from research on the big ideas of expectation, variability, and distribution as a guide to instruction. Investigate how students are reasoning about them, then provide tasks that will challenge how they are currently reasoning so that they grow in their understanding along the levels in the research frameworks. Get students in the habit of posing their own statistical questions and using the statistical investigation cycle from GAISE to explore and answer their statistical questions. Make sure that attention to variability is foremost throughout the statistical investigation cycle. Most of all, empower students to be competent and confident with the big concepts and processes of statistics, and with the nature of statistical argumentation.

References

Ben-Zvi, D. (2004) Reasoning about variability in comparing distributions. *Statistics Education Research Journal, 3*(2), 42-63.

Ben-Zvi, D. & Garfield, J. (Eds.). (2004). *The Challenge of Developing Statistical Literacy, Reasoning, & Thinking.* Dordrecht, The Netherlands: Kluwer.

Biehler, R., Frischmeier, D., Reading, C., & Shaughnessy, J. M. (2018). Reasoning about data. In D. Ben-Zvi, K. Makar, &, J. Garfield (Eds.), *International Handbook of Research in Statistics Education* (pp. 139-192). Cham, Switzerland: Springer International.

Case, C. & Jacobbe, T. (2018). *Statistics Education Research Journal*, *17*(2), 9-29. Retrieved from https://iase-web.org/documents/SERJ/SERJ17(2)_Case.pdf

Cobb, G. (2007). The introductory statistics course. A Ptolemaic curriculum? *Technology Innovations in Statistics Education,* 1(1), Article 1. Retrieved from http://escholarship.org/uc/item/6hb3k0nz

Cobb, G. & Moore, D. (1997). Mathematics, statistics, and teaching. *The American Mathematical Monthly,104*(9), 801-823.

Franklin, C., Kader, G., Mewborne, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2007). *Guidelines for Assessment and Instruction in Statistics Education.* Alexandria, VA: ASA.

Franklin, C.A., Barbabliotti, A. E., Case, C. A., Kader, G. D., Scheaffer, R. L., & Spangler, D. A. (2015). *The Statistical Education of Teachers.*

Ganadesikan, M., Scheaffer, R. L., Landwehr, J. M., Watkins, A. E., Barbella, P. Kepner, J., Newman, C. M., Obremski, T. E., & Swift, J. (1995). *Quantitative Literacy Series.* New York: Pearson Learning (Dale Seymour Publications).

Garfield, J., & Ben-Zvi. D. (2008). *Developing Students' Statistical Reasoning. Connecting Research and Practice.* Springer.com.

Kader. G., & Jacobbe, T. (2013). *Developing Essential Understanding of Statistics Grades 6-8.*
Reston, VA: NCTM

Konold, C., & Pollatsek, A. (2002). Data analysis as a search for signals in noisy processes. *Journal for Research in Mathematics Education, 33*(4), 259-289.

Langrall, C. W., Makar, K., Nilsson, P., & Shaughnessy, J. M. (2017). Teaching and learning probability and statistics: An integrated perspective. In J. Cai (Ed.), *Compendium for Research in Mathematics Education.* Reston, VA: NCTM.

Makar, K., & Rubin. A. (2018). Learning about statistical inference. In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.). *International Handbook of Research in Statistics Education* (pp. 261-294). Cham, Switzerland: Springer International.

Melitou, M. (2002). Conceptions of variation. A literature review. Statisics Education Research Journal, *1*(1), 46-52.

Mokros, J., & Russell, S. J. (1995). Children's concepts of average and representativeness. *Journal for Research in Mathematics Education, 26*(1), 20-39.

Moore, D. (1997). New pedagogy and new content: The case of statistics. *International Statistics Review, 65*(2), 123-165.

Mosteller, F., Kruskal, W. H., Link, R. F., Pieters, R. S., & Rising, G. R. (1973). *Statistics by Example.* Reading, MA: Addison-Wesley.

National Council of Teachers of Mathematics. (1980). *An Agenda for Action.* Reston, VA: NCTM.

National Council of Teachers of Mathematics. (1989). *Curriculum and Evaluation Standards for School Mathematics.* Reston, VA: NCTM.

National Council of Teachers of Mathematics. (2000). *Principles and Standards for school Mathematics.* Reston, VA: NCTM.

Noll, J., & Shaughnessy J. M. (2012). Aspects of students' reasoning about variation in empirical

sampling distributions. *Journal for Research in Mathematics Education, 43*(5), 509-556.
Peck, R., Gould, R., & Miller, S. J. (2013). *Developing Essential Understanding of Statistics Grades 9-12*. Reston, VA: NCTM.
Reading C., & Reid, J. (2006). An emerging hierarchy of reasoning about distribution: From a variation perspective. *Statistics Education Research Journal, 5*(2), 46-68.
Reid, J., & Reading, C. (2008). Measuring the development of students' consideration of variation. *Statistics Education Research Journal, 7*(1), 40-59.
Reading, C., & Shaughnessy, J. M. (2004). Reasoning 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.
Rossman, A. (2008). Reasoning about informal statistical inference: One statistician's point of view. *Statistics Education Research Journal, 7*(2), 5-19.
Shaughnessy, J. M. (1992). Research in probability and statistics: Reflections and directions. In D. Grouws (Ed.). *Handbook of Research on Mathematics Teaching and Learning* (pp. 465-494). New York, NY: Macmillan.
Shaughnessy, J. M. (2007). Research on statistics learning and reasoning. In F. Lester (Ed.). *Second Handbook of Research on Mathematics Teaching and Learning* (pp. 957-1009). Reston, VA: NCTM.
Shaughnessy, J. M., Garfield, J., & Greer, B. (1996). Data handling. In A. J. Bishop, K. Clements, C. Keitel, J. Kilpatrick, & C. Laborde (Eds.). *International Handbook of Mathematics Education* (pp. 205-237). Dordrecht, The Netherlands: Kluwer.
Shaughnessy, J. M., Watson, J., Reading, C. & Moritz, J. (1999. April). School mathematics students' Acknowledgement of statistical variation: There's more to life than centers. Paper presented at the research pre-session of the 77th annual meeting of the National Council of Teachers of Mathematics, San Francisco, CA.
Tanur, J., Mosteller, F., Kruskal, W. H., Link, Pieters, R. S., & Rising, G. R. (1972). *Statistics: A Guide to the Unknown.* San Francisco, CA: Holden-Day.
Toruk, R., & Watson, J. M. (2000). Development of the concept of statistical variation: An exploratory study. *Mathematics Education Research Journal. 9*, 60-82.
Watson, J. M., (2009). The influence of variation and expectation on developing awareness of distribution. *Statistics Education Research Journal, 8*(1), 32-61.
Watson, J, M., & Kelly, B. A. (2004). Expectation versus variation: Students' decision making in a chance environment. *Canadian Journal of Science, Mathematics, and Technology Education, 4,* 371-396.
Watson, J. M., & Moritz, J. B. (2000). The longitudinal development of understanding average. *Mathematical Thinking and Learning, 2,* 11-50.
Watson, J. M., Callingham, R. A., & Kelly, B. A. (2007). Students appreciation of expectation and variation as a foundation for statistical inference. *Mathematical Thinking and Learning, 9*(2), 83 – 130.
Watson, J. M., Kelly, B. A., Callingham, R. A, & Shaughnessy, J. M. (2003). The measurement of school Students' understanding of statistical variation. *International Journal of Mathematical Education in Science and Technology, 34*(1), 1-29.
Watson, J. M., Fitzallen, N., & Carter, P. (2013). *Top Drawer Teachers: Statistics.* Adelaide, Australia: Australian Association of Mathematics Teachers and Services Australia. Retrieved from http://topdrawer.aamt.edu.au/Statistics
Zeiffler, A., Garfield, J., Del Mas, R., & Reading, C. (2008). A framework to support research on informal inferential reasoning. *Statistics Education Research Journal, 7*(2), 40-58.
Zeiffler, A., Garfield, J., & Fry, E. (2018). What is statistics education? In D. Ben-Zvi, K. Makar, and J. Garfield (Eds.). *International Handbook of Research in Statistics Education,* 37-70.