Chapter 6
Building Concept Images of Fundamental Ideas in Statistics: The Role of Technology

Gail Burrill

Abstract Having a coherent mental structure for a concept is necessary for students to make sense of and use the concept in appropriate and meaningful ways. Dynamically linked documents based on TI© Nspire technology can provide students with opportunities to build such mental structures by taking meaningful statistical actions, identifying the consequences, and reflecting on those consequences, with appropriate instructional guidance. The collection of carefully sequenced documents is based on research about student misconceptions and challenges in learning statistics. Initial analysis of data from preservice elementary teachers in an introductory statistics course highlights their progress in using the documents to cope with variability in a variety of contextual situations.

Keywords Concept image · Deviation · Distribution
Interactive dynamic visualization · Mean · Variability

6.1 Introduction

Educators have suggested that visual images provide an important tool for learning (e.g. Breen 1997). Dreyfus (1991) argued that the "status of visualization in mathematics education should and can be upgraded from that of a helpful learning aid to that of a fully recognized tool for learning and proof" (vol. I: p. 33). Presmeg (1994) suggested that visualizing mathematical concepts is a means to develop understanding. Interactive dynamic technology can be an important factor in helping students build these images. This view is supported by a number of studies that suggest strategic use of technological tools can help students transfer mental images of concepts to visual interactive representations that lead to a better and more robust understanding of the concept (e.g. Artigue 2002; Guin and Trouche 1999). In particular, technology plays a central role in teaching and learning statistics, perhaps a greater role

G. Burrill (✉)
Program in Mathematics Education, Michigan State University, East Lansing, MI, USA
e-mail: burrill@msu.edu

G. Burrill and D. Ben-Zvi (eds.), *Topics and Trends in Current Statistics Education Research*, ICME-13 Monographs, https://doi.org/10.1007/978-3-030-03472-6_6

than for many other disciplines (Chance et al. 2007). A variety of researchers have investigated the role of technology in the learning of statistics (c.f., Ben-Zvi 2000; Burrill 2014; Watson and Fitzallen 2016). In many classrooms, however, the use of technology can too easily focus only on organizing data, graphing and calculating. The perspective throughout this chapter is that technology, particularly interactive dynamic technology, can and should be used for more than "doing the work", a view supported by the American Statistical Association's Guidelines for Assessment and Instruction in Statistics Education (GAISE) that stresses the use of technology for developing conceptual understanding as well as carrying out analyses (Franklin et al. 2007).

6.2 The Potential of Interactive Dynamic Technology

Content specific learning technologies provide many opportunities for developing understanding of statistical concepts. Interactive dynamic technology allows students to link multiple representations—visual, symbolic, numeric and verbal—and to connect these representations to support understanding (Sacristan et al. 2010; Biehler et al. 2013; Burrill 2014). For example, a regression line can be dynamically linked to a visualization of the residual squares and the numerical sum of the squared residuals. Such interactive linking, where one object is manipulated and all related representations are instantly updated, supports investigations into varying assumptions and asking "what if" questions that can lead to making and testing conjectures and result in a better understanding of the concepts involved (Ben-Zvi 2000). Computer simulation activities enable students to experience variability by comparing random samples, generating simulated distributions of sample statistics, and observing the effect of sample size on sampling distributions (delMas et al. 1999; Hodgson 1996). The ability to display multiple screens simultaneously allows students to contrast different graphs of the same data or notice how changing a data point affects a distribution. Spreadsheet features provide opportunities for managing large sets of data, enabling students to investigate subsets of the data for similarities and differences, for example, sorting a data set according to gender to compare curfews or spending money. Many misconceptions held by students about statistical concepts can be confronted using technology in a "predict and-test" strategy, establishing a cognitive dissonance that can help students change their thinking about a concept (e.g., Posner et al. 1982). Students can predict what they think they will observe (e.g., expected shape of a distribution) and then use the technology to obtain immediate feedback on their thinking.

From a more fundamental perspective, however, interactive dynamic technology has the potential to help students create robust *concept images* of key statistical ideas, a necessary step in being able to fluently and effectively reason with and apply those ideas (Oehrtman 2008). A concept image can be described as the total cognitive structure including the mental pictures and processes associated with a concept built up in students' minds through different experiences associated with the ideas (Tall and Vinner 1981). Without a coherent mental structure, students are left to construct an understanding based on ill formed and often misguided connections and images (Oehrtman 2008). The work of understanding subsequent topics is then built on isolated understandings specific to each topic (e.g., center as separate from spread, distribution as a set of individual outcomes, randomness as accidental or unusual). This makes it difficult for students to see and work with the relationship among the images needed for deep understanding, for example, to understand the distinction among the distribution of a population, the distribution of a sample from that population, and the distribution of a statistic computed from samples from the population.

A *concept definition* can be thought of as the words used to specify the concept, which are typically related in some way to a student's personal experiences with the concept. As students engage in new experiences related to the concept, a student's concept image changes and evolves into a personal concept definition. For example, a student's first image of mean might be "add and divide"—an image of the specific rule for calculating the mean of a set of data. If the concept image of mean remains at this level, students will struggle when they are asked to interpret a mean in context or approximate a mean from the graph of a distribution. The educational goal should be to provide students with experiences that will help them move to a more formal understanding of the concept, supported by the development of rich interconnected concept images/definitions, that is accepted by the community at large (Tall and Vinner 1981).

Piaget argued that an individual's conceptual structure is based on the actions or the coordination of actions on physical or mental objects made by the individual (Piaget 1970, 1985). Given this stance, instruction beginning with formal definitions would seem to be contrary to the direction in which abstraction occurs. Oehrtman (2008) suggests three important features of instructional activities compatible with Piaget's theory of abstraction. First, the underlying structure that is the target for student learning should be reflected in the actions they do. Because these actions come before conceptual understanding, they should be stated in terms accessible to students rather than formal definitions, enabling students' eventual concept images to build from conceptual structures that make sense to them because of their previous actions. Second, students' actions should be repeated and organized with provisions for feedback and ways to respond to this feedback. And third, students should use these actions in structurally similar problems in a variety of contexts to develop a robust abstraction of the concept.

This chapter describes a sequence of applet-like documents, *Building Concepts: Statistics and Probability* (2016) *(BCSP)* developed according to an "action/consequence principle" aligned with Oehrtman's features (2008). The materials were created to exploit the affordances of an interactive dynamic environment in developing robust conceptual structures for key statistical concepts, designed for teaching introductory statistical concepts.

6.3 Building Concepts

6.3.1 An Action/Consequence Principle

To make sense of ideas, it is necessary to have appropriate conceptual structures, and it is impossible to communicate about concepts without any representations (Bakker and van Eerde 2014). In his semiotic theory on the use of diagrams as ways to represent relationships, Peirce describes *diagrammatic reasoning* as that which involves constructing a diagram, experimenting with it, and reflecting upon the results. He emphasizes that in the experimenting state, "thinking in general terms is not enough. It is necessary that something should be DONE. In geometry, subsidiary lines are drawn. In algebra, permissible transformations are made. Thereupon the faculty of observation is called into play" (CP 4.233—CP refers to Peirce's collected papers, volume 4, section 233). Because the learner has done something with the diagram or representation, he is forced to consider the consequences of the action from a different perspective than that originally in his mind (Peirce 1932, 1.324).

In *Building Concepts*, these three steps are embodied in an "action/consequence" principle, where the learner can "deliberately take a mathematical action, observing the consequences, and reflecting on the mathematical implications of the consequences" (Mathematics Education of Teachers II 2012, p. 34). In statistics, the actions might involve grouping data points in a certain way, changing bin widths in histograms, moving data points, generating random samples from a population, changing the sample size, or moving a line. The consequences might be different visual representations of the data, changes in numerical summaries, noting what remains constant and what changes with the action, or a shift in patterns. By reflecting on the changes they see in response to statistically meaningful actions, students are engaged in actively processing, applying, and discussing information in a variety of ways (National Research Council 1999; Michael and Modell 2003) and can begin to formulate their own concept images and conceptual structures of key statistical ideas.

From another perspective, the theories of Mezirow (1997), Kolb's learning cycle model (1984), and the work of Zull (2002) on brain theory all suggest that people learn through the mechanism of participating in an immersive mathematics experience, reflecting on these experiences, and attempting similar strategies on their own. Mezirow introduced the notion of transformative learning as a change process that transforms frames of reference for the learner. Key elements in this process are an "activating event" (Cranton 2002) that contributes to a readiness to change (Taylor 2007). This is followed by critical reflection where the learner works through his understanding in light of the new experiences, considering the sources and underlying premises (Cranton 2002). The third element of this process is reflective discourse or dialogue in an environment that is accepting of diverse perspectives (Mezirow 2000). The final step is acting on the new perspective, central for the transformation to occur (Baumgartner 2001). These four elements elaborate on Kolb's early model of experiential learning (1984) as a cycle containing four parts: concrete experi-

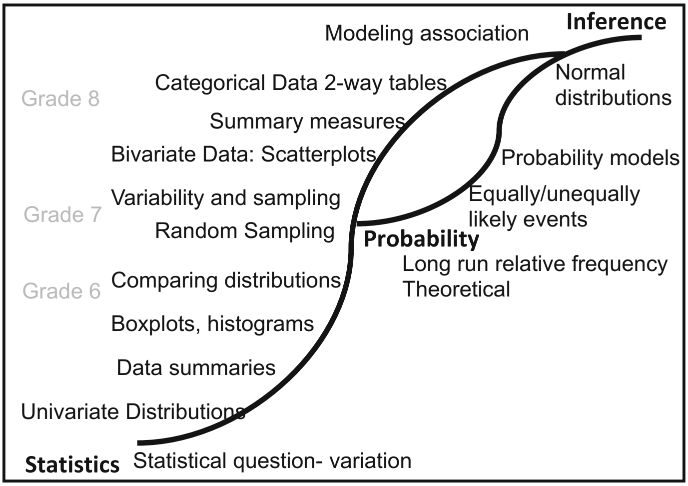


Fig. 6.1 Statistics and probability—a coherent progression

ence, reflective observation, abstract conceptualization, and active experimentation; experimentation leads once again to concrete experience. This cycle, informed by Oehrtman's key features, is embodied in the action/consequence principle underlying the activities in *Building Concepts*.

6.3.2 Content Framework

The content in the *BCSP* activities is based on the Common Core State Standards (CCSS) progressions documents (2011), narratives describing the learning progression of a topic based on the research on cognitive development and on the logical structure of mathematics/statistics. Taken as a whole, the activities and corresponding dynamic files cover the key concepts typically in introductory school statistics (Fig. 6.1). Static pictures or examples contained in the progression document are made interactive in the activities. In addition, the activities have been designed in light of the research related to student learning, challenges and misconceptions.

6.3.3 The Activities

The core of the activities are applet-like, dynamic interactive files, not intended to be used for "doing" statistical procedures but rather to provide a mental structure for reasoning about statistical concepts that can support the transition to procedural fluency. When students have a solid conceptual foundation, they can engage in statistical thinking, are less susceptible to common errors, less prone to forgetting and are able to see connections and build relationships among ideas (NRC 1999).

6.3.3.1 Framing of Tasks

The tasks in each activity focus on using the interactive documents to create experiences that can contribute to the development of a particular statistical concept. They were constructed following the advice of Black and Wiliam (1998) with respect to formative assessment: "Tasks have to be justified in terms of the learning aims that they serve, and they can work well only if opportunities for pupils to communicate their evolving understanding are built into the planning (p. 143)." Thompson (2002) argued that the goal of a task is to have students participating in conversations that foster reflection on some mathematical "thing". Thus, the majority of tasks in the activities create opportunities to discuss particular statistical objects or ideas that need to be understood and to ensure that specific conceptual issues and misconceptions will arise for students as they engage in discussions.

6.3.3.2 Misconceptions

The tasks in the activities have been designed in light of the research related to student learning, challenges and misconceptions (e.g., Zehavi and Mann 2003). For example, a common misconception in statistics relates to boxplots: the longer one of the four sections in the plot, the more data in that section (Bakker et al. 2005). To build a mental image of the connection between the data and a boxplot, in the interactive file a dot plot "morphs" into the boxplot, and students can compare the number of data values in each section of the boxplot (Fig. 6.2). Moving points in the dot plot immediately displays the effect on the corresponding boxplot (Fig. 6.3), reinforcing the fact that medians and quartiles are summary measures based on counting.

The activity *Equally Likely Events* was designed specifically to address the misconception that every outcome has a 50% chance of occurring (Fischbein et al. 1991). In this activity, students generate a distribution of the eleven possible sums of the faces when two dice are tossed and compare the distribution to a distribution of the outcomes of spinning a spinner divided into eleven equal regions. The visualization of the distributions as the number of repetitions is increased makes explicit how a random sample reflects the characteristics of the population. In *Comparing Distributions*, students explicitly contrast histograms and bar graphs to confront the confusion they often have distinguishing between the two representations. They consider the limitations of bar graphs in understanding the story of the typical income, education and life expectancy in various regions of the world and compare what is lost or gained when the data are represented in boxplots, histograms, or dot plots. Students create histograms with a large amount of variability and with little variability to challenge

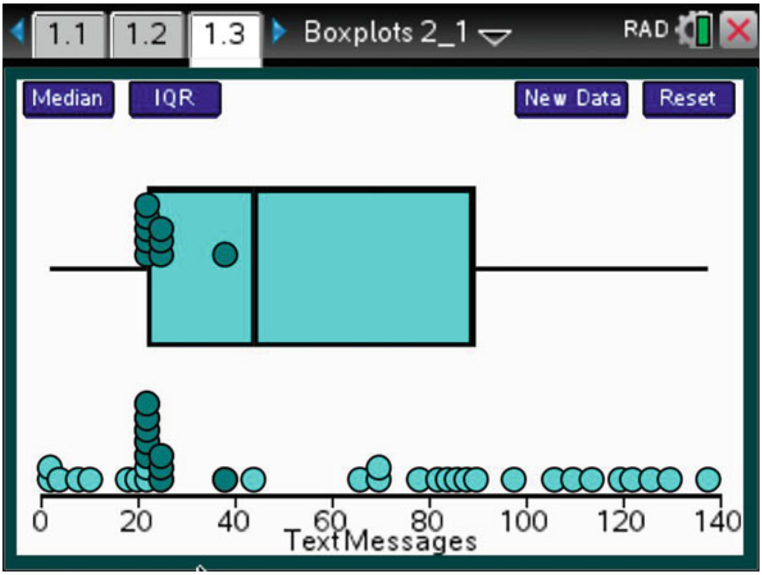


Fig. 6.2 Connecting boxplots to data

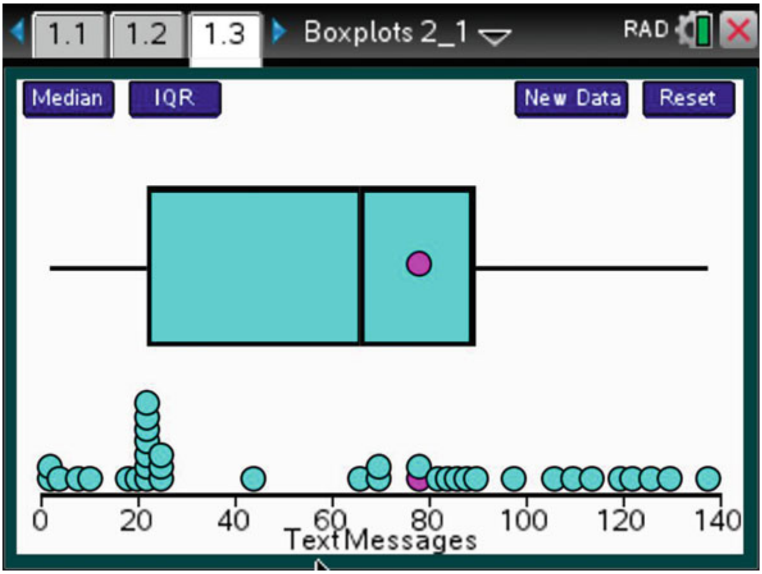


Fig. 6.3 Moving a data point

the misconception that variability is defined by the range or by a peak rather than the spread around the mean (delMas and Liu 2005; Matthews and Clark 2003).

6.3.3.3 Posing Questions

In addition to making sure that the tasks surface misconceptions and develop understanding of "tough to teach/tough to learn" concepts, the questions for each of the activities were created using the general guidelines below:

1. Activate prerequisite knowledge before it is used; e.g., "Remember the importance of thinking about shape, center and spread when talking about distributions of data. Describe the distribution on page 1.3." (*Introduction to Histograms*)
2. Point out things to notice so students focus on what is important to observe; e.g., "Select Sample. Describe the difference between the points on the number line at the top left and the point on the number line at the right." (*Sample Means*)
3. Ask for justifications and explanations; e.g., "Make a conjecture about which data set will have the largest mean. Explain why you think your conjecture might be correct. Use the file to check your thinking." (*Mean as Balance Point*)
4. Make connections to earlier tasks or to an immediately previous action taken by the student (questions should not come out of the blue); e.g., "Return to your answers for question 2 and see if you want to change them now that you have looked at the values when they are ordered." (*Median and Interquartile Range*)
5. Include both positive and negative examples in developing understanding of definitions, theorems and rules; e.g., "Which of the following are true? Give an example from the Ti-Nspired file to support your reasoning. (a) The smallest and largest values of any distribution are outliers. (b) Not all distributions have outliers. (c) An outlier will be more than one boxplot width plus half of the width of the boxplot to the left and right of the box. d) The segments on each side of the box always extend 1½ IQRs beyond the LQ and the UQ." (*Outliers*)
6. Have students consider the advantages/disadvantages of each approach when it is possible to carry out a task using multiple strategies; e.g., "Which, if any, of the three estimation methods—educated guess, judgment sample, or random sample—do you think is more likely to give a sample that is most representative of the population? Why?" (*Random Samples*)
7. Be explicit about possible misconceptions: e.g., "Work with a partner to create two reasonable distributions for the number of pairs of shoes owned by the students in a class, either by moving or adding points, to get (1) a distribution with little variability in the number of pairs of shoes owned by most of the class, and (2) a distribution where there is a lot of variability in the number of pairs of shoes owned by the class. Choose a bin width that seems best for your distribution. Describe your distribution (shape, center and spread). Explain why you think one of your distributions has very little variability and the other has a lot of variability." (*Introduction to Histograms*)

The next section provides several examples of using the dynamic interactive files to develop concept images related to core statistical ideas. These include distributions, measures of center and spread, and random behavior including sampling variability.

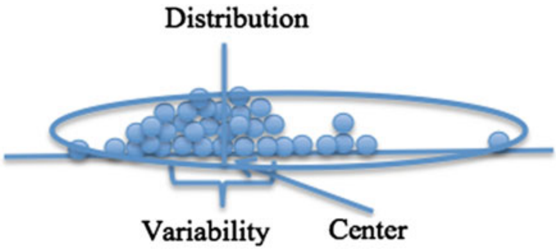


Fig. 6.4 Image of distribution, center and variability

6.4 Developing Concepts

6.4.1 Distributions

A statistical distribution might be defined as "an arrangement of values of a variable showing their observed or theoretical frequency of occurrence" (The Free Dictionary). Wild (2006) suggests, however, that "… the notion of "distribution" is, at its most basic, intuitive level, "the pattern of variation in a variable," (p. 11) and further notes that focusing on what a distribution is will not be as productive as focusing on helping students build a mental image of how data can be distributed. According to Wild, because distributions are such a fundamental component of statistical reasoning the goal should be on how you can reason with distributions and not on how do you reason about distributions.

A well-documented problem observed by statistics educators is that students tend to perceive data as a series of individual cases and not as a whole that has characteristics and properties not observable in any of the individual cases (i.e., Bakker and Gravemeijer 2004; Ben-Zvi and Arcavi 2001; Hancock et al. 1992). They suggest that students need to develop a conceptual structure in which data sets are thought of as aggregates where the concept image of how data can be distributed includes features related to shape, center and variability around the center (Fig. 6.4).

In the first *BCSP* activity, *Introduction to Data*, students investigate numerical lists and dot plots of the maximum recorded speeds and life spans of different animal types (Fig. 6.5) with the goal of building a mental image of a distribution of the data. Students begin by identifying individual animals or data points (How fast is a tiger?). They are then asked to talk about the distribution as a whole, connecting words such as "clumps", "bumps", "piles", "dots are spread out" (Bakker and Gravemeijer 2004; Cobb et al. 2003) to shapes, eventually building images of distributions that can be described using language accepted in the statistical community: mound shaped, symmetric, skewed, uniform. Transitions from language such as "all bunched at one end" to "skewed" or "the dots are spread out" to "the spread is large" are important steps in the formation of concepts (Peirce 1998). The interactive files allow students to notice how changing a data point affects a distribution and to experiment with removing data points in a distribution to see the effect on the shape (How will the

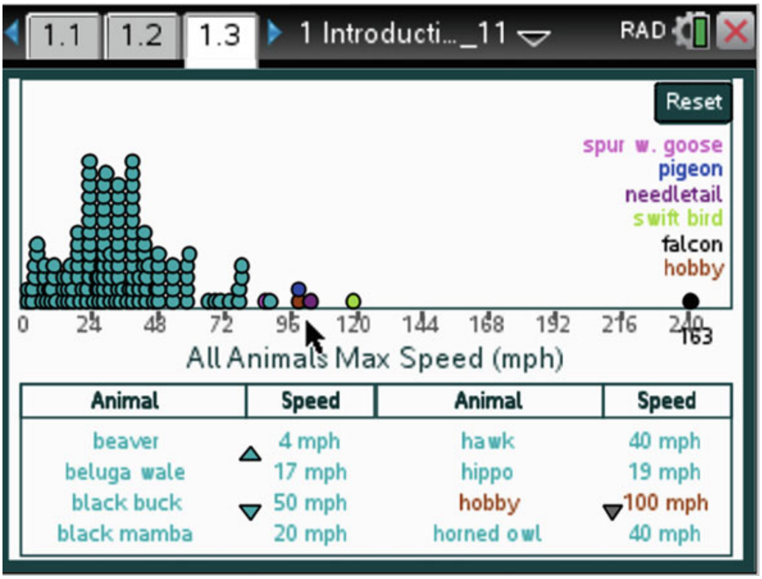
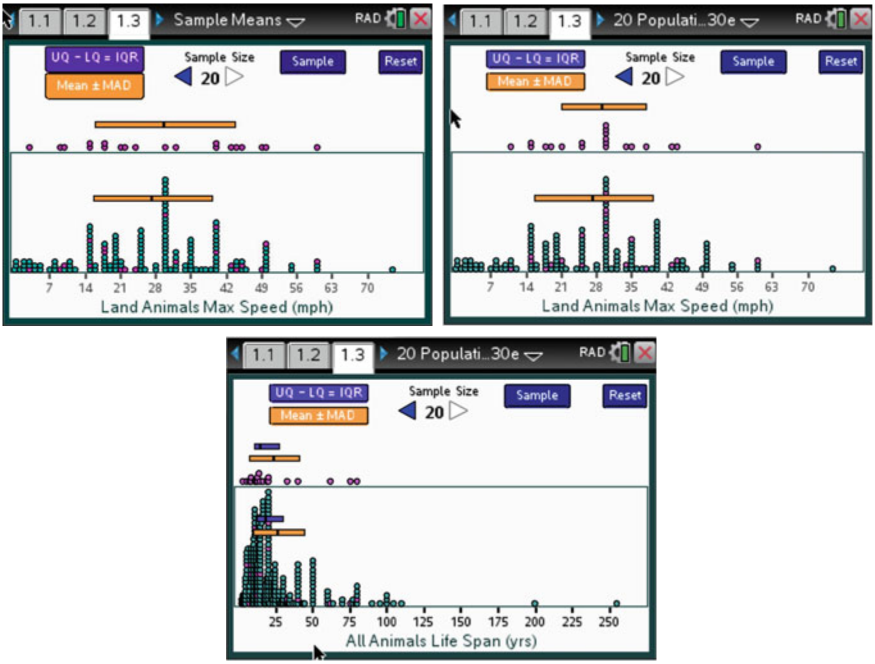


Fig. 6.5 Maximum speeds of types of animals

shape change if the maximum speed of the Peregrine falcon is removed from the distribution? If the speeds for all of the birds are removed?).

Building from the conceptual structures students have formed in this initial work and mindful of the principle that students should use these actions in structurally similar problems in a variety of contexts, the concept of distribution is revisited in other activities, such as those which develop the connection between measures of center and spread and the shape of a distribution. The technology allows students to cycle through a variety of data sets, providing opportunities to recognize distributions when the mean may not represent the largest cluster of data points, and the median may be a more useful measure of center.

The concept of distribution is revisited again in the context of sampling. The concept structures students have developed for reasoning with distributions are extended to consider distributions of a sample from a population, where, for example, they examine the distributions of maximum recorded speeds for a sample of animal types, the plot on the top in Figs. 6.6, 6.7, and 6.8, and distinguish this from the distribution of the maximum recorded speeds for all of the animal types, the plot on the bottom in Figs. 6.6, 6.7, and 6.8. Repeatedly taking samples provides contexts in which the distribution of sample maximum recorded speeds reflects the population but varies from sample to sample as do the summary measures (mean ± mean absolute deviation or median and interquartile range) associated with the random samples of the maximum speeds, reflected by the horizontal bars in Figs. 6.6, 6.7, and 6.8.



Figs. 6.6–6.8 Populations, samples and sampling variability

6.4.2 Mean and Standard Deviation

Students' concept images of measures of center and spread seem to be fragile. Misconceptions or superficial understanding of measures of center have been well documented (Friel 1998; Groth and Bergner 2006; Mokros and Russell 1995; Watson and Moritz 2000). Students often can perform the computations but cannot apply or interpret the concepts in different situations and have correspondingly ill formed notions of variability. In the past many texts introduced the mean and median as measures of center in a single lesson, and several lessons later or in another chapter, if at all, introduced measures of variability. Treating center and spread together supports the creation of a mental structure of the notion that measures of spread are connected to "spread around what"—some value indicating a measure of center (see Fig. 6.4); deviations are measures of distance from the mean, and the interquartile range (IQR) is a measure of the distance between the first and third quartiles and thus around the median. Experiences with these different interpretations of center and variability, can help students build a mental structure mindful of the need to take both measures into account when reasoning about variation in a variety of situations (Shaughnessy et al. 1999) and can help them recognize that either measure alone tells an incomplete story about the context.

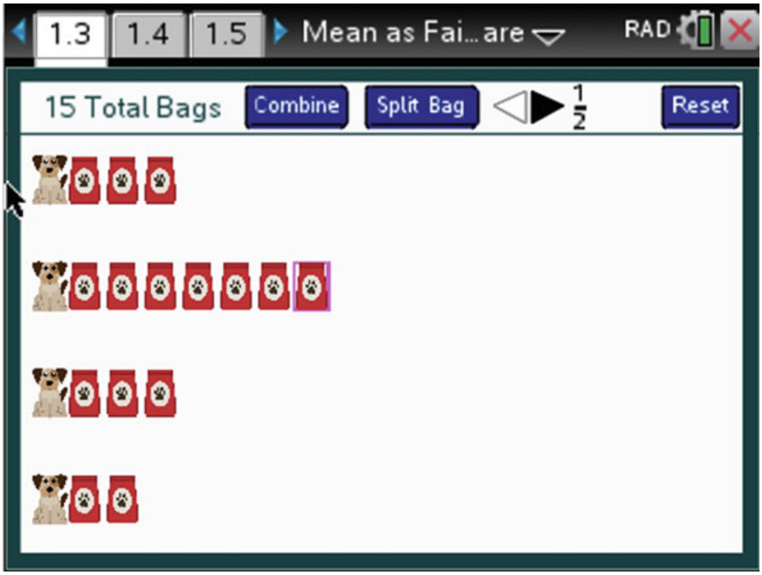


Fig. 6.9 Mean as leveling

In *Building Concepts*, median and interquartile range are introduced in one activity followed by activities related to mean and mean absolute deviation (which is introduced as a precursor to standard deviation). The literature suggests that typically students have problems interpreting the mean and applying it appropriately (e.g., Garfield and Ben-Zvi 2005). To counter this, the activities explicitly develop the concept of mean as "fair share" in two ways. The activities endeavor to build mental images of (1) fair share as "leveling off" where students drag dog food bags from the dogs who have the most bags to dogs with fewer bags (Fig. 6.9) until all of the dogs have the same number of bags; and (2) fair share as pooling, where all of the contributors (the dogs) put their bags of dog food into a group (Fig. 6.10), and the entire group is then divided equally among the total number of contributors (dogs) (Fig. 6.11). Both approaches contribute to developing images needed for complete understanding of mean; the first develops an understanding of how to interpret the mean as a measure of center, and the later leads directly to the procedural algorithm typically used to compute a mean.

Recognizing the difficulty students have shifting their images of bar graphs as ways to describe distributions of data to graphs involving quantitative data displayed on a number line, one file focuses on connecting numerical (16 total bags) and pictorial representations to a dot plot, where students observe how the dot plot changes as the pictorial representations are moved (Fig. 6.12). The fact that all of the dots are in a vertical line at four when each dog has four bags of dog food lays the groundwork for considering the mean as a balance point.

Fig. 6.10 Pooling

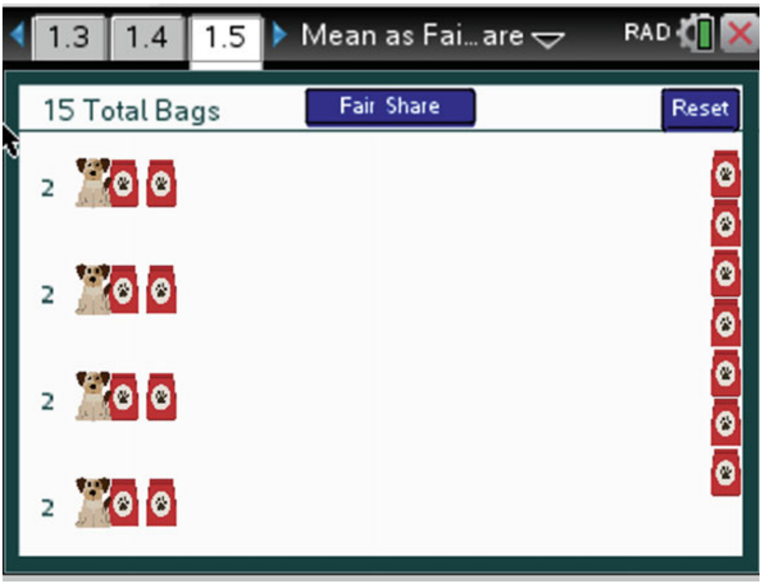


Fig. 6.11 Dividing up the pool

Mean As Balance Point is set in the context of soccer tournaments, where the task is to distribute a given total number of goals in a tournament to achieve a mean number

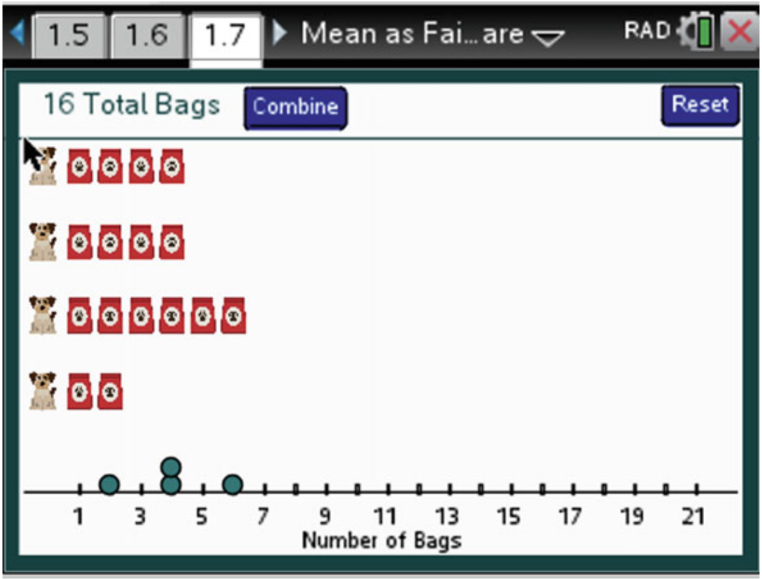


Fig. 6.12 Connecting representations

of six goals for the nine teams involved (Kader and Mamer 2008). An important part of the reflection step in the action/consequence principle is for students to describe what they see in the diagrams (Figs. 6.13 and 6.14) then, with support, learn to abstract from the picture the notion of deviation, where deviation in itself can have characteristics (Pierce 1998). The goal of this activity is to give students experience in describing deviations, resulting in the development of an image for the concept of deviation as an object itself and to eventually link deviation to the concept of variability. Students move dots representing the number of goals scored by a soccer team to "balance" the dots on the number line, given that the mean number of goals for all of the teams has to be six, and certain constraints must be satisfied (e.g., no teams scored six goals, two teams scored two goals, one team scored three goals, and one team scored nine goals). They can notice how changing a data point affects the distribution of goals and explore how the "deviations" from the mean are related to whether the segment containing the distribution of goals is balanced.

Students identify a measure to rank different tournaments (distributions) in terms of "most evenly matched teams" with the assumption that, in a tournament with perfectly matched teams, every team scores the same number of goals (Fig. 6.15). This leads to the mean absolute deviation as a measure of spread around the mean and the notion of mean as balance (Fig. 6.16). The development "uncouples" the words "standard" and "deviation", giving students the opportunity to build an image of the word deviation in a simple context before they think about standard deviations.

Associating shapes with measures of center and variability can help students develop an understanding of what these measures mean graphically and numerically

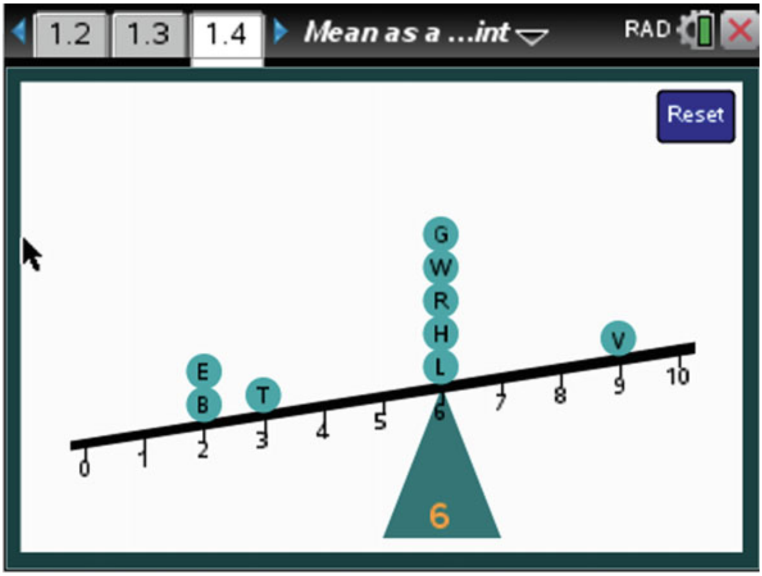


Fig. 6.13 Goals in a tournament

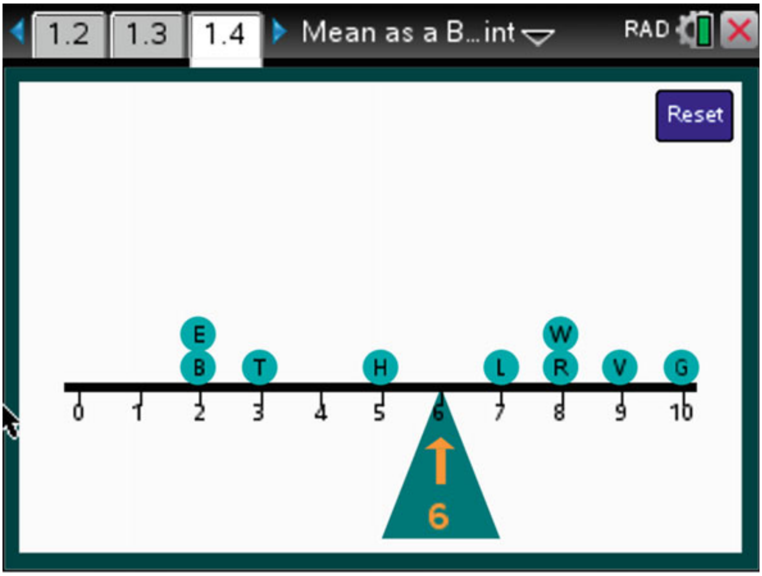


Fig. 6.14 Constraints satisfied

(Garfield and Ben-Zvi 2005). Connecting the image of deviation to a mental image of variation around the mean, students use the technology to make conjectures about

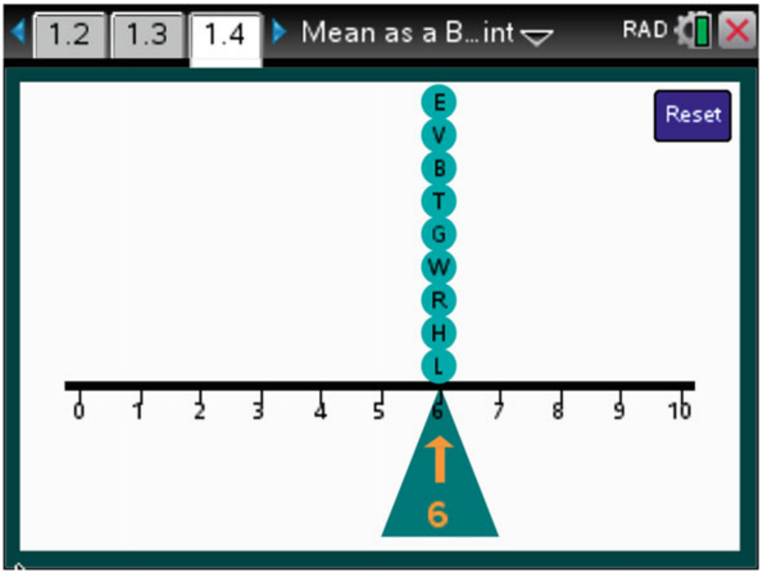


Fig. 6.15 Evenly matched

Fig. 6.16 Ranking soccer tournaments in terms of the most "evenly balanced" teams

the measures of center and spread for randomly generated distributions of scores and can instantly check their conjectures (Fig. 6.17). The technology supports students in continuing to build their mental images by making visible the connections among numerical, visual and algebraic representations as they interpret data in a table and relate the data and summary measures to a graph (Figs. 6.18 and 6.19).

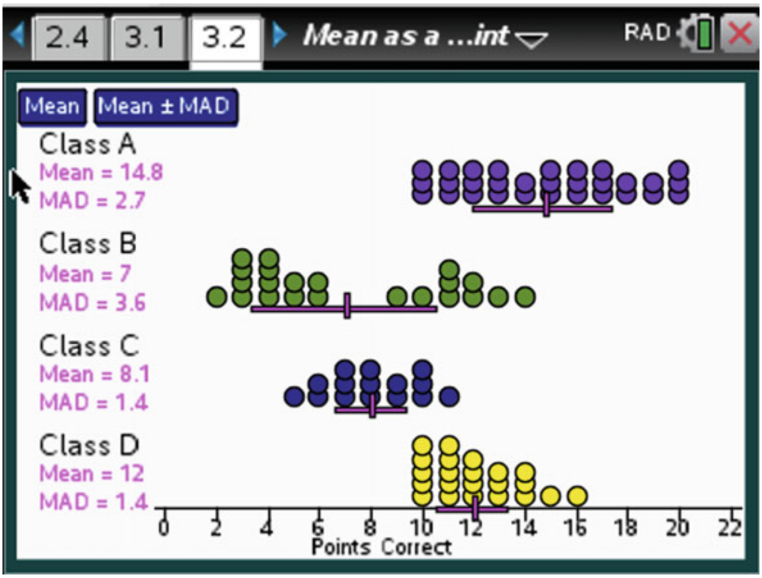


Fig. 6.17 Checking conjectures

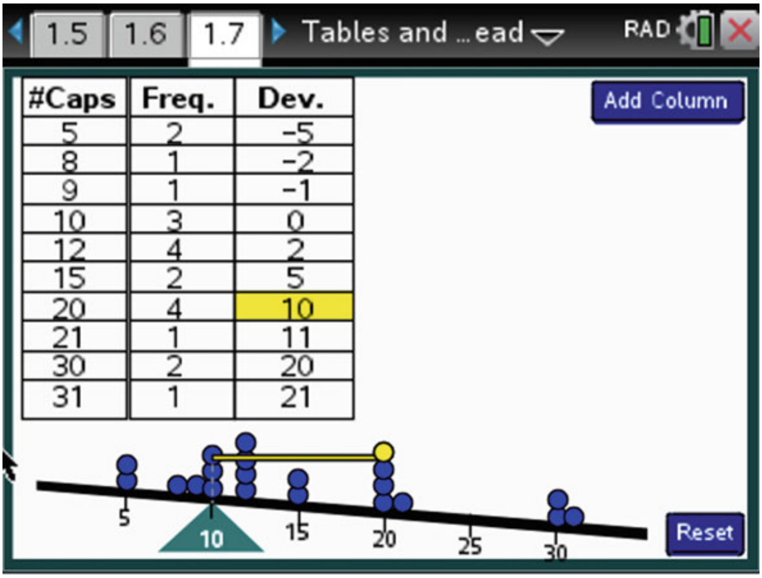


Fig. 6.18 Deviations

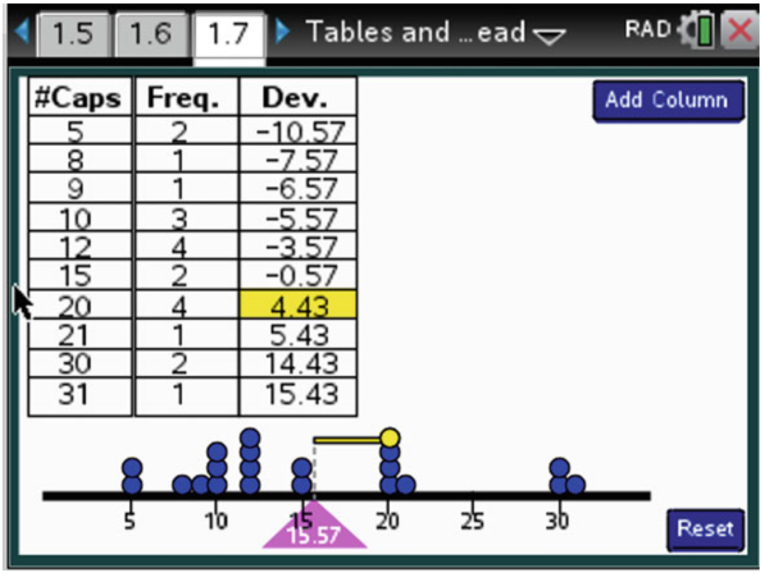


Fig. 6.19 Deviations from the mean

6.4.3 Random Behavior

To most people, "random" events in their lives can be those that are surprising, due to luck or fate, not repeatable or happen just due to "chance" (Batanero 2015). Thus, the natural language learners bring to developing a concept image of randomness is often in conflict with the formal concept definition itself. This can seriously impede the learning of a formal notion of randomness. Students having such a potential conflict in their concept image may be comfortable with their own interpretations of randomness and simply regard the formal theory as not realistic and superfluous (Tall and Vinner 1981). Furthermore, students are bothered by the notion of predicting with some certainty the behavior of a distribution but being unable to predict a specific outcome (Konold 1989). Some believe that it is not possible to apply mathematical methods (statistics) to study random phenomena, because of their unpredictability. Some also believe they can predict or control the outcomes in a random process (Langer 1975).

Given the complexity of building concept images that will enable students to confront their intuitive notions about a random event and align them with the meaning used in statistics, the learning experiences in which students engage need to be carefully designed. Batanero (2015) recommends one possible sequence. First, students learn to discriminate certain, possible and impossible events in different contexts, using the language of chance, and compare an analysis of the structure of an experiment with the frequencies of data collected from repeated experiments to estimate probability. In a second stage, students should move to the study of materials lacking symmetry properties (e.g., spinners with unequal areas, thumbtacks), where they can only estimate probability from frequencies. The next stage is to investigate real data available from the media, Internet, government or other sources (e.g., sports, demographic, or social phenomena). Finally, students simulate simple situations where the essential features of the situation are modeled by the model used in the simulation and irrelevant properties are disregarded.

Aligned with this framework, the *BCPS* activities introduce the notion of probability using a game where students choose which of two options (i.e., odd, prime) is more likely to occur in drawing ten cards each with a number from one to 10. They have opportunities to play the same game several times and to figure out strategies for winning (the number of successes over the total number of outcomes), giving them experiences that can lead to the creation of a mental structure for estimating probabilities when it is possible to list the outcomes. The technology allows students to simulate the probability using the relative frequencies of a long sequence of drawing cards. The next step is to contrast this situation, where the theoretical outcomes are clear, to a situation where nothing is known about the probability of an outcome (getting a blue chip in drawing a chip from a bag with an unknown number of white and blue chips), using long run relative frequencies to estimate the probability of an outcome (i.e., blue chip). Students generate many repetitions of the experiment, formulate questions or predictions about the trend in the outcomes, collect and analyze data to test their conjectures, and justify their conclusions based on these data. This approach allows students to visualize randomness as a dynamic process in contrast to a printed copy of a random sequence that seems to lose the essence of what random means (Johnston-Wilder and Pratt 2007). The typical sequence of results obtained through repetition lacks a pattern (Fig. 6.20) at the onset. However, "In this apparent disorder, a multitude of global regularities can be discovered, the most obvious being the stabilization of the relative frequencies of each possible result" (Batanero 2015) (see Fig. 6.21).

Students learn that streaks and clusters can appear in a sequence of random outcomes. Technology can be used to create situations involving a cognitive dissonance to help students change their ways of thinking about the concept. In the activity *Choosing Random Samples* students draw names from a hat to identify four students out of 28 to hand in their homework on a given day (Fig. 6.22) (supporting Oehrtman's third feature of instructional activities (2008)—experience the concept in a variety of situations). Students believe that random behavior somehow balances out in the short run, and once you have been selected you will no longer be called on (Fischbein and Schnarch 1997; Jones et al. 2007; Konold 1989). Simulation allows the process to be repeated many times, and students soon recognize that by chance, a random selection will typically have several students chosen two or even three times in a five-day week. Simulation in a real context can help students establish a better understanding of the nature of randomness. This pattern of random behavior, in the short-term unpredictable but in the long-term stable, is revisited in generating distributions of sample statistics. For random samples selected from a population, students can observe that medians and means computed from random samples will vary from

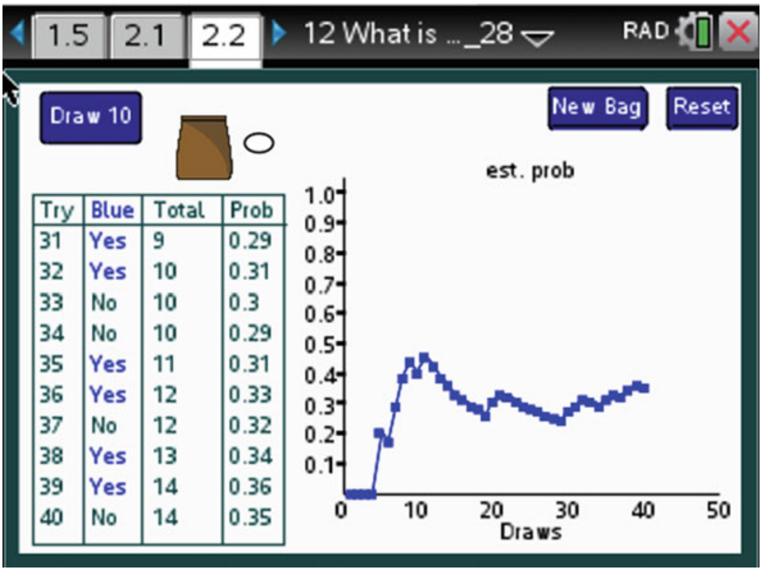


Fig. 6.20 Initial variability

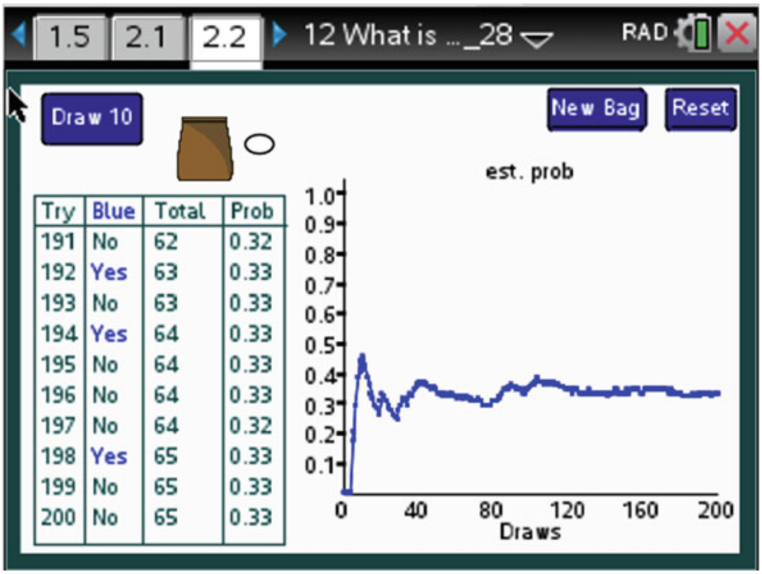


Fig. 6.21 Relative frequency stabilizes

sample to sample and that making informed decisions based on such sample statistics requires some knowledge of the amount of variation to expect.

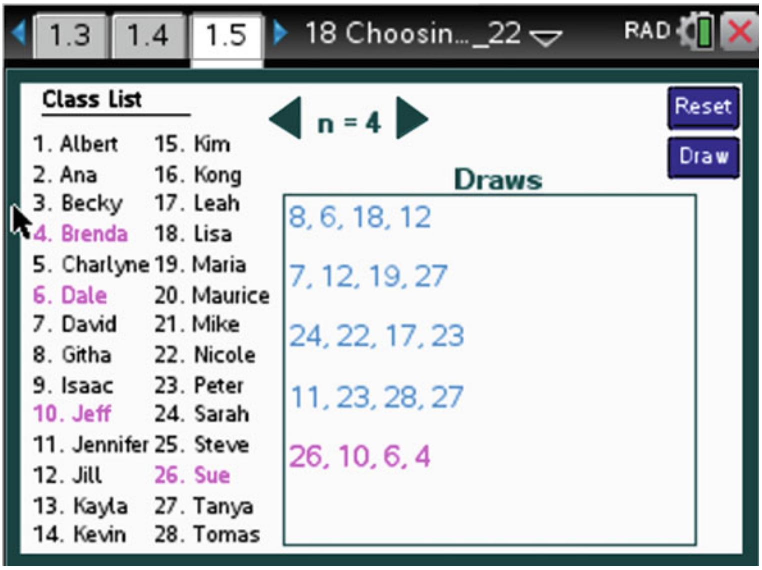


Fig. 6.22 Randomly chosen frequency stabilizes

6.4.4 Sampling Distributions

Students often confuse the three types of distributions related to sampling: distribution of a population, the distribution of a sample from that population, and the sampling distribution of a sample statistic (Wild 2006). In *Samples and Proportions* the notion of distribution is extended from considering the distribution of a population itself and the distribution of a sample from that population to a third kind of distribution, a sampling distribution of statistics calculated from the samples taken from the population (Fig. 6.23).

Students generate many different simulated sampling distributions for a given sample size of the proportion of females from a population that is known to be half female. They discover that each of these distributions seems be mound shaped and symmetric, centered on the expected value with a consistent range for the number of females in the sample over repeated simulations. A subtle but critical point for learners is that a shift from counts to proportions allows comparison of distributions with different sample sizes and opens up opportunities to think about what is invariant and what is not as the sample size changes and why. The mental image here is highly dependent on noticing that the distinguishing feature is the labeling of the axes. Students can observe that for a sample of a given size, the simulated distribution of the number of females in a sample from a population that is 30% female visually overlaps with the sampling distribution of the number of females in a sample from a population that is 50% female (Fig. 6.24). This leads informally to the concept of margin of error.

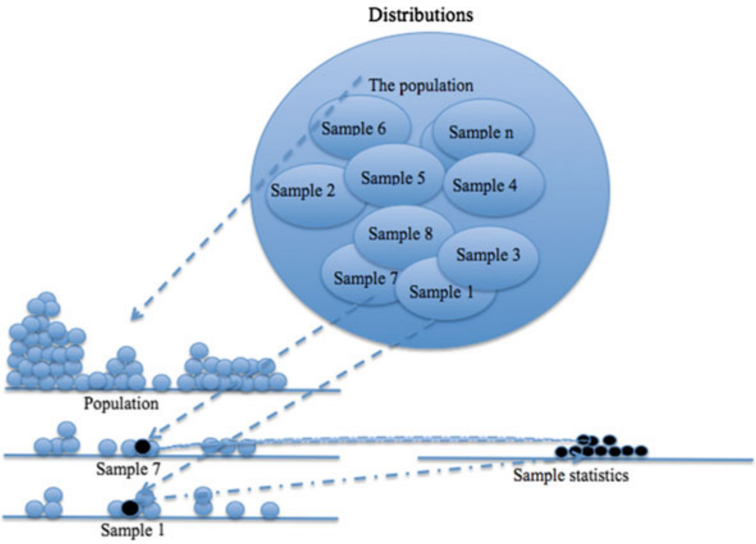


Fig. 6.23 Concept images of three related but distinct notions of distribution

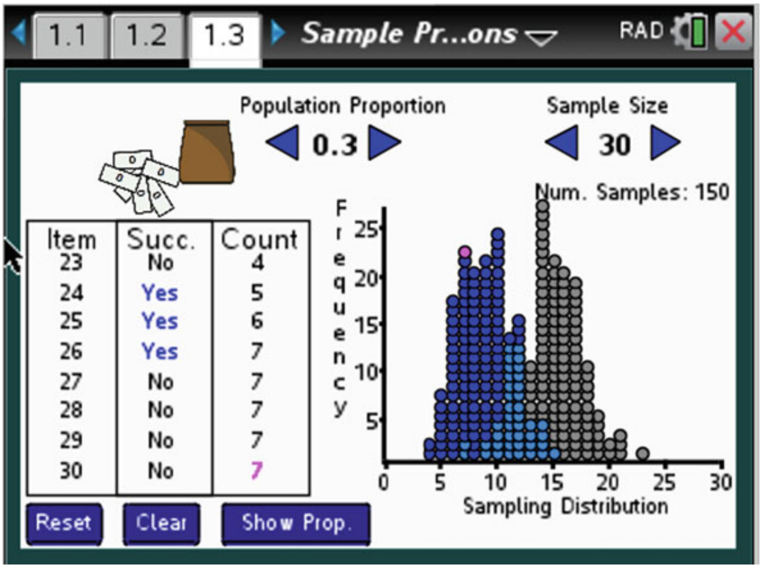


Fig. 6.24 Comparing simulated distributions from two populations

The discussion above described several of the 24 different activities, each addressing particular concepts typically in an introductory statistics course at the school level and as outlined in the CCSS progressions for Statistics and Probability (2011). The files are accompanied by supporting materials that include (1) a description of the statistical thinking that underlies the file; (2) a description of the file and how to use it; (3) possible mathematical objectives for student learning; (4) sample questions for student investigation; and (5) a set of typical assessment tasks. The activities have been developed for use on a TI© Nspire platform (iPad app, computer software or handheld) and can be downloaded at no cost from the *Building Concepts* website (https://education.ti.com/en/building-concepts/activities/statistics).

6.5 Implementation

The interactive documents were used in a semester long statistics course for elementary preservice students. Students had their own computers, and they accessed the materials using the TI Nspire software, although they did use other statistical software packages towards the end of the course. The goals of the course were to enable students to be literate consumers of statistical data related to education and to give them tools and strategies for their own teaching. Student learning experiences were designed with attention to the action/consequence cycle described in Sect. 3.1, an action or activating event, critical reflection, reflective discourse and taking actions based on the new perspective. The students typically worked in pairs or groups on predesigned tasks using the technology to investigate situations, make and test conjectures, usually comparing their results with classmates and engaging in student led discussions on their thinking about the ideas.

6.5.1 Background

The students were sophomores or juniors in the elementary teacher preparation program at a large Midwestern university. They all had selected a mathematics emphasis for their certification (and had taken calculus, which enabled them to interpret the point of inflection on a relatively normal distribution as approximately one standard deviation from the mean); 24 had no prior experience with statistics; three had taken an Advanced Placement statistics course in high school and two had taken a university statistics course.

In keeping with the GAISE framework (Franklin et al. 2007) and the focus of the research, one emphasis in the course was on variability. The next section briefly describes how the interactive documents and action/consequence cycle played out with respect to helping students understand the role of variability in statistical reasoning.

6.5.2 Instruction

An “activating event” was an activity or question that engaged students’ curiosity and lead to an investigation of a statistical concept. In the second week of the course, students were asked: How long did it typically take a student in our class to get to campus today? Students made conjectures, then lined up across the classroom according to their times (without talking to make things interesting). The class reflected on the visual representation they had formed; eventually realizing they needed to regroup as they had neglected to consider scale and had just ordered themselves. The distribution of their times now had several clusters and one clear outlier at 90 min. The distribution was reproduced on the board, and the class considered the question: How would you describe the “typical time”. This led to language like “a center cluster”, which motivated a discussion of median, interquartile range and how these ideas would be useful in identifying the typical time to campus. (It turned out to be from 5 to 15 min for half of the students with the median at 8 min.) The outlier was described as surprising.

Cycling through the process, students applied the questions “what is typical” and “what would be surprising” in a variety of situations and new experiences, working with their classmates in randomly assigned groups on tasks creating different graphical representations (action/consequence) and considering the variability in each (reflection). As in the literature, they initially confused variability with range: “[A has] Most in variability: there are many observations on pairs of shoes that people own covering a wide range.” But the majority of students were able to correctly make statements such as: “In variability the height of the peaks don’t matter. Additionally, we are only really looking at the center and how the graph looks around it.”

In the application part of the cycle, students began to use the concept of variability in meaningful ways. For example, looking at the achievement of fourth grade students in science, one student wrote, “An interesting thing about scores from 2015, as can be seen in Fig. 6.2, is that there is an outlier, a state with an average score of 140. This is interesting because while this score is an outlier in the 2015 data, this is not an outlier in 2005, in fact it is part of the lower quartile range. This indicates that in 2015 a score that low would be somewhat unusual, because higher scores are being achieved in science by all of the other states.” They did however continue to struggle with language: “When comparing the western states’ funding to the eastern state’s [sec] funding, the eastern states have a larger range in terms of IQR.”

In a similar fashion, activities such as the soccer tournament described above motivated the use of mean $\pm$ MAD (mean absolute deviation) and eventually the standard deviation. Students used simulation to establish what is typically the pattern for a sampling distribution for a given population proportion and sample size. Individually repeating the process over and over (action) and comparing distributions across the class (consequence) gave students the opportunity for critical reflection and to recognize the distribution will always be mound shaped and symmetric with the mean and median around the expected value and one standard deviation at approximately the point of inflection if a smooth curve were drawn over the simulated sampling distribution. Students noticed the variability in number of successes for a large number of samples of the same size is typically bounded as they simulated the event many times; for example, for a population proportion of 0.5 and sample size 100, the number of successes will rarely be less than 35 or more than 65. "Is it surprising?" led to the activating question—just what does it mean to be surprising? What if for the example, an observed outcome was 34. How do we communicate the notion of surprise at such an observation to other classmates? The discussion and reflection on how to quantify or find a measure for surprising led to the notion of significance.

6.5.3 *Initial Results*

An initial analysis of some of the data suggests that students for the most part have a relatively solid grasp of variability. For example, the variability around student scores on a state achievement test was given as margins of error (Fig. 6.25). When asked on the final exam, for which student, A, B, or C was the margin of error most problematic, 48% were able to correctly identify student B and 28% answered choice A or C with appropriate reasoning.

Some had the correct answer but incorrect or unclear reasoning; e.g., "This is because with the margin of error, there are lower possible answers than the other students that students with that scale score could have obtained."

In comparing the standard deviations for the length of time males and females could stand on one foot, 48% of the students associated the standard deviation with the mean but they continued to struggle with precision of language ("The difference in standard deviation between males and females indicates that females are more clustered around the mean."), while 45% described the variability in general terms without reference to the mean. When asked what image comes to mind when you think about variability, 21% of the students connected variability to the spread

Fig. 6.25 Student achievement results

around the mean or median, 17% gave a general description such as "apartness", "differentness", while 31% gave a measure (MAD, IQR, standard deviation).

In comparing the achievement scores of boys and girls, a response such as the following was typical: "…we find that 64.4% of the time this could occur just by chance. We can use this to answer our question and say that grade 5 girls weren't more likely to score below basic than grade 5 boys because our results could have occurred just by chance. They're not statistically significant so we can't say that either gender is more likely to do worse than the other gender based on these results."

6.6 Conclusions, Future Directions and Research Recommendations

The study was purely observational, with no comparison group or controls for factors such as prior knowledge (although the class as a whole came with little exposure to statistics), which limits any conclusions that can be made. Initial results do seem to suggest the approach has potential for supporting the development of student understanding of variability in multiple statistical contexts. However, the research connecting concept images to visualization to dynamic interactive technology is sparse and a space where much work remains. Some possible questions include:

- What aspects of pedagogy are significant in the use of visualization through dynamic interactive technology in learning mathematics?
- How can teachers help learners use dynamic interactive technology to make connections between visual and symbolic representations of statistical ideas?
- How might dynamic interactive technology be harnessed to promote statistical abstraction and generalization?
- How do visual aspects of interactive dynamic technology change the dynamics of the learning of statistics?

In 1997, Ben-Zvi and Friedlander noted that technology for teaching and learning has evolved over the years, progressively allowing the work to shift to a higher cognitive level enabling a focus on planning and anticipating results rather than on carrying out procedures. Since then technology has provided powerful new ways to assist students in exploring and thinking about statistical ideas, allowing students to focus on interpretation of results and understanding concepts rather than on computational mechanics. While visualizing mathematical concepts has been considered important in developing understanding of these concepts, dynamic interactive technology provides opportunities for students to build more robust conceptual images—to develop video images in their minds as they consider what a concept means in a given context. The *Building Concepts* work thus far suggests that interactive dynamic technology affords students opportunities to build concept images of statistical concepts that align with desirable conceptions of those concepts. The carefully designed action/consequence documents seem to have the potential to be useful tools in providing students with the experiences they need to develop the robust concept images of core statistical concepts that will enable them to use statistics as a way to reason and make decisions in the face of uncertainty.

References

Artigue, M. (2002). Learning mathematics in a CAS environment: The genesis of a reflection about instrumentation and the dialectics between technical and conceptual work. *International Journal of Computers for Mathematical Learning, 7*(3), 245–274.

Bakker, A., Biehler, R., & Konold, C. (2005). Should young students learn about boxplots? In G. Burrill & M. Camden (Eds.), *Curriculum development in statistics education: International association for statistics education 2004 roundtable* (pp. 163–173). Voorburg, the Netherlands: International Statistics Institute.

Bakker, A., & Gravemeijer, K. (2004). Learning to reason about distributions. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 147–168). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Bakker, A., & Van Eerde, H. (2014). An introduction to design-based research with an example from statistics education. In A. Bikner-Ahsbahs, C. Knipping, & N. Presmeg (Eds.), *Doing qualitative research: Methodology and methods in mathematics education* (pp. 429–466). New York: Springer.

Batanero, C. (2015). Understanding randomness: Challenges for research and teaching. In K. Kriner (Ed.), *Proceedings of the ninth congress of the European Society for Research in Mathematics Education* (pp. 34–49).

Baumgartner, L. M. (2001). An update on transformational learning. In S. B. Merriam (Ed.), *New directions for adult and continuing education, no. 89* (pp. 15–24). San Francisco, CA: Jossey-Bass.

Ben-Zvi, D. (2000). Toward understanding the role of technological tools in statistical learning. *Mathematical Thinking and Learning, 2*(1–2), 127–155.

Ben-Zvi, D., & Arcavi, A. (2001). Junior high school students' construction of global views of data and data representations. *Educational Studies in Mathematics, 45*(1–3), 35–65.

Ben-Zvi, D., & Friedlander, A. (1997). Statistical thinking in a technological environment. In J. B. Garfield & G. Burrill (Eds.), *Research on the role of technology in teaching and learning statistics* (pp. 45–55). Voorburg, The Netherlands: International Statistical Institute.

Biehler, R., Ben-Zvi, D., Bakker, A., & Makar, K. (2013). Technology for enhancing statistical reasoning at the school level. In M. A. Clements, A. Bishop, C. Keitel, J. Kilpatrick & F. Leung (Eds.), *Third international handbook of mathematics education* (pp. 643–690). Springer.

Building Concepts: Statistics and Probability. (2016). Texas Instruments Education Technology. http://education.ti.com/en/us/home.

Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80*(2), 139–144.

Breen, C. (1997). Exploring imagery in P, M and E. In E. Pehkonen (Ed.), *Proceedings of the 21st PME International Conference, 2*, 97–104.

Burrill, G. (2014). Tools for learning statistics: Fundamental ideas in statistics and the role of technology. In *Mit Werkzeugen Mathematik und Stochastik lernen*[Using Tools for Learning Mathematics and Statistics], (pp. 153–162). Springer Fachmedien Wiesbaden.

Chance, B., Ben-Zvi, D., Garfield, J., & Medina, E. (2007). The role of technology in improving student learning of statistics. *Technology Innovations in Statistics Education*, *1* (2). http://repositories.cdlib.org/uclastat/cts/tise/vol1/iss1/art2.

Cobb, P., McClain, K., & Gravemeijer, K. (2003). Learning about statistical covariation. *Cognition and Instruction, 21*(1), 1–78.

Common Core State Standards. (2010). *College and career standards for mathematics*. Council of Chief State School Officers (CCSSO) and National Governor's Association (NGA).
Cranton, P. (2002). Teaching for transformation. In J. M. Ross-Gordon (Ed.), *New directions for adult and continuing education, no. 93* (pp. 63–71). San Francisco, CA: Jossey-Bass.
delMas, R., Garfield, J., & Chance, B. (1999). A model of classroom research in action: Developing simulation activities to improve students' statistical reasoning. *Journal of Statistics Education*, [Online] *7*(3). (www.amstat.org/publications/jse/secure/v7n3/delmas.cfm).
delMas, R., & Liu, Y. (2005). Exploring students' conceptions of the standard deviation. *Statistics Education Research Journal*, *4*(1), 55–82.
Dreyfus, T. (1991). On the status of visual reasoning in mathematics and mathematics education. In F. Furinghetti (Ed.), *Proceedings of the 15th PME International Conference*, *1*, 33–48.
Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., et al. (2007). *Guidelines for assessment and instruction in statistics education (GAISE) report: A Pre-K–12 curriculum framework*. Alexandria, VA: American Statistical Association.
Fischbein, E., Nello, M. S., & Marino, M. S. (1991). Factors affecting probabilistic judgments in children and adolescents. *Educational Studies in Mathematics, 22*(6), 523–549.
Fischbein, E., & Schnarch, D. (1997). The evolution with age of probabilistic, intuitively based misconceptions. *Journal for Research in Mathematics Education, 28*(1), 96–105.
Free Dictionary. http://www.thefreedictionary.com/statistical+distribution.
Friel, S. (1998). Teaching statistics: What's average? In L. J. Morrow (Ed.), *The teaching and learning of algorithms in school mathematics* (pp. 208–217). Reston, VA: National Council of Teachers of Mathematics.
Garfield, J., & Ben-Zvi, D. (2005). A framework for teaching and assessing reasoning about variability. *Statistics Education Research Journal, 4*(1), 92–99.
Groth, R., & Bergner, J. (2006). Preservice elementary teachers' conceptual and procedural knowledge of mean, median, and mode. *Mathematical Thinking and Learning, 8*(1), 37–63.
Gould, R. (2011). *Statistics and the modern student*. Department of statistics papers. Department of Statistics, University of California Los Angeles.
Guin, D., & Trouche, L. (1999). The complex process of converting tools into mathematical instruments: The case of calculators. *International Journal of Computers for Mathematical Learning, 3*(3), 195–227.
Hancock, C., Kaput, J., & Goldsmith, L. (1992). Authentic inquiry with data: Critical barriers to classroom implementation. *Educational Psychologist, 27*(3), 337–364.
Hodgson, T. (1996). The effects of hands-on activities on students' understanding of selected statistical concepts. In E. Jakbowski, D. Watkins & H. Biske (Eds.), *Proceedings of the eighteenth annual meeting of the North American chapter of the international group for the psychology of mathematics education* (pp. 241–246).
Johnston-Wilder, P., & Pratt, D. (2007). Developing stochastic thinking. In R. Biehler, M. Meletiou, M. Ottaviani & D. Pratt (Eds.), *A working group report of CERME 5* (pp. 742–751).
Jones, G., Langrall, C., & Mooney, E. (2007). Research in probability: Responding to classroom realities. In F. K. Lester (Ed.), *The second handbook of research on mathematics* (pp. 909–956). Reston, VA: National Council of Teachers of Mathematics (NCTM).
Kader, G., & Mamer, J. (2008). Contemporary curricular issues: Statistics in the middle school: Understanding center and spread. *Mathematics Teaching in the Middle School, 14*(1), 38–43.
Kolb, D. (1984). *Experiential learning: Experience as the source of learning and development*. New Jersey: Prentice-Hall.
Konold, C. (1989). Informal conceptions of probability. *Cognition and Instruction, 6*(1), 59–98.
Langer, E. J. (1975). The illusion of control. *Journal of Personality and Social Psychology, 32*(2), 311–328.
Learning Progressions for the Common Core Standards in Mathematics: 6–8 Progression probability and statistics (Draft). (2011). Common Core State Standards Writing Team.

Mathematics Education of Teachers II. (2012). Conference Board of the Mathematical Sciences. Providence, RI and Washington, DC: American Mathematical Society and Mathematical Association of America.

Mathews, D., & Clark, J. (2003). *Successful students' conceptions of mean, standard deviation and the central limit theorem*. Unpublished paper.

Mezirow, J. (1997). Transformative learning: Theory to practice. In P. Cranton (Ed.), *New directions for adult and continuing education, no. 74.* (pp. 5–12). San Francisco, CA: Jossey-Bass.

Mezirow, J. (2000). Learning to think like an adult: Core concepts of transformation theory. In J. Mezirow & Associates (Eds.), *Learning as transformation: Critical perspectives on a theory in progress* (pp. 3–34). San Francisco, CA: Jossey-Bass.

Michael, J., & Modell, H. (2003). *Active learning in secondary and college science classrooms: A working model of helping the learner to learn*. Mahwah, NJ: Erlbaum.

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

National Research Council. (1999). In J. Bransford, A. Brown & R. Cocking (Eds.), *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academy Press.

Oehrtman, M. (2008). Layers of abstraction: Theory and design for the instruction of limit concepts. In M. Carlson & C. Rasmussen (Eds.), *Making the connection: Research and teaching in undergraduate mathematics education* (pp. 1–21).

Peirce, C. S. (1932). In C. Hartshorne & P. Weiss (Eds.), *Collected papers of Charles Sanders Peirce 1931–1958*. Cambridge, MA: Harvard University Press.

Peirce, C. S. (1998). *The essential Peirce: Selected philosophical writings, Vol. 2 (1893–1913).* The Peirce Edition Project. Bloomington, Indiana: Indiana University Press.

Piaget, J. (1970). *Structuralism*, (C. Maschler, Trans.). New York: Basic Books, Inc.

Piaget, J. (1985). *The equilibration of cognitive structures* (T. Brown & K. J. Thampy, Trans.). Chicago: The University of Chicago Press.

Posner, G., Strike, K., Hewson, P., & Gertzog, W. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education, 66*(2), 211–227.

Presmeg, N. C. (1994). The role of visually mediated processes in classroom mathematics. *Zentralblatt für Didaktik der Mathematik: International Reviews on Mathematics Education, 26*(4), 114–117.

Sacristan, A., Calder, N., Rojano, T., Santos-Trigo, M., Friedlander, A., & Meissner, H. (2010). The influence and shaping of digital technologies on the learning— and learning trajectories—of mathematical concepts. In C. Hoyles & J. Lagrange (Eds.), *Mathematics education and technology—Rethinking the mathematics education and technology—Rethinking the terrain: The 17th ICMI Study* (pp. 179–226). New York, NY: Springer.

Shaughnessy, J., Watson, J., Moritz, J., & Reading, C. (1999). School mathematics students' acknowledgement of statistical variation. In C. Maher (Chair), *There's more to life than centers*. *Presession Research Symposium, 77th Annual National Council of Teachers of Mathematics Conference*, San Francisco, CA.

Tall, D., & Vinner, S. (1981). Concept image and concept definition in mathematics with particular reference to limits and continuity. *Educational Studies in Mathematics, 12*(2), 151–169.

Taylor, E. W. (2007). An update of transformative learning theory: A critical review of the empirical research (1999–2005). *International Journal of Lifelong Education, 26*(2), 173–191.

Thompson, P. (2002). Didactic objects and didactic models in radical constructivism. In K. Gravemeijer, R. Lehrer, B. V. Oers & L. Verschaffel (Eds.), *Symbolizing, modeling and tool use in mathematics education* (pp. 191–212). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Watson, J., & Fitzallen, N. (2016). Statistical software and mathematics education: Affordances for learning. In L. D. English & D. Kirshner (Eds.), *Handbook of international research in mathematics education* (3rd ed., pp. 563–594). New York, NY: Routledge.

Watson, J., & Moritz, J. (2000). Developing concepts of sampling. *Journal for Research in Mathematics Education, 31*(1), 44–70.

Wild, C. (2006). The concept of distribution. *Statistics Education Research Journal,* 5(2), 10–26. http://www.stat.auckland.ac.nz/serj.

Zehavi, N., & Mann, G. (2003). Task design in a CAS environment: Introducing (in)equations. In J. Fey, A. Couco, C. Kieran, L. McCullin, & R. Zbiek (Eds.), *Computer algebra systems in secondary school mathematics education* (pp. 173–191). Reston, VA: National Council of Teachers of Mathematics.

Zull, J. (2002). *The art of changing the brain: Enriching the practice of teaching by exploring the biology of learning*. Alexandria VA: Association for Supervision and Curriculum Development.