UNDERSTANDING SAMPLING DISTRIBUTIONS: THE ROLE OF INTERACTIVE DYNAMIC TECHNOLOGY

Gail Burrill
Michigan State University, East Lansing MI, USA
Burrill@msu.edu

Interactive dynamic technology can help students build concept images of core statistical ideas, in particular, ideas related to distributions and sampling, enabling them to make inferences from data and in the process help address common misconceptions about these ideas. Applet-like documents allow students to take meaningful statistical actions in diverse situations, immediately see the consequences, and reflect on those consequences in terms of specified learning outcomes. Initial results of the analysis of preservice elementary teachers thinking after using the documents based on a framework adapted from the Solo taxonomy are promising.

INTRODUCTION

Conceptual knowledge can be thought of as an implicit or explicit understanding of the principles that govern a content area and how the units of knowledge in that area are connected (Rittle-Johnson, Siegler, & Alibali, 2001). “A fundamental premise of virtually all student development models is that true intellectual growth will not occur unless the foundational skills (conceptual understanding) is firmly grounded” (Felder & Brent, 2005, p. 57). Tall and Vinner (1981) describe this understanding as a concept image: 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. 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).

As students engage in new experiences related to a concept, a student’s concept image changes and evolves. For example, a student’s first image of the mean might be an image of a formula used for computing a mean. If the concept image remains at this level, students will struggle when they are asked to use and interpret means in different situations such as developing the concept of standard deviation. 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 & Vinner, 1981).

BACKGROUND

A number of studies suggest the strategic use of technological tools can help students transfer mental images of concepts to visual interactive representations leading to a better and more robust understanding of the concept (e.g. Guin & Trouche, 1999). Interactive dynamic technology can provide students with “live” visualizations of a concept that have the potential to enable students to build robust images of the properties, processes and relationships connected to the concept (Drijvers, 2015). Interactive dynamic technology allows students to link multiple representations of problem scenarios − visual, symbolic, numeric and verbal – and to connect these representations to support understanding (Sacristan et al., 2010). For example, a regression line can be dynamically linked to a visualization of residual squares and the sum of the squared residuals. Computer simulations allow students to build representations of variability by generating simulated distributions of sample statistics, comparing random samples, and observing the effect of sample size on sampling distributions (delMas, Garfield, & Chance, 1999). Interacting with dynamic files allows students to build a concept image that includes “movie clips” of the features of the concept that can become the basis for understanding. Applet-like electronic documents can leverage dynamic linkages between representations to address specific learning outcomes for statistical concepts. This paper describes the use of such applet-like documents based on materials from *Building Concepts: Statistics and*

In: S Budgett (Ed.), *Decision Making Based on Data*
Proceedings of the Satellite conference of the International Association for Statistical Education (IASE),
August 2019, Kuala Lumpur, Malaysia. ©2019 ISI/IASE iase-web.org/Conference_Proceedings.php

Probability (2015) to help students develop robust conceptual structures related to sampling distributions.

Oehrtman (2008) suggests three important features of instructional activities related to the development of concept images. First, the underlying structure that is the target for student learning should be reflected in the actions they do. Second, students' actions should be repeated and organized with provisions for feedback and ways to respond to this feedback. And third, students should repeat these actions in structurally similar problems in a variety of contexts to develop a robust abstraction of the concept. The *Building Concepts* apps were designed with an "action-consequence-reflection" principle that reflects Oehrtman's thinking, where the learner can deliberately take an action, observe the consequences, and reflect on the statistical implications of the consequences. In statistics, the actions might involve grouping data points in a certain way or changing the sample size. The consequences might be different visual representations of the data, changes in numerical summaries, or noting what remains constant and what changes with the action. 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) and can begin to formulate their own concept images and conceptual structures of key statistical ideas.

Typically, many statistical concepts are intertwined and together provide the foundation for statistical reasoning and decision making, always in the presence of variability. For example, distrbutional reasoning begins with reasoning about a distribution of data, understanding different measures of center and variability and how they relate to the shape of a distribution; a sound understanding of these ideas is foundational for reasoning about sampling distributions (Chance, delMas & Garfield, 2004). Reading and Reid (2006) describe a heirarchical framework for distribution using two cycles: the first, key elements of distribution (center, variability, density, skewness, and outliers) and the second, using distribution for statistical inference. This paper postulates another cycle in the heirarchy, reasoning about randomness and simulated sampling distributions, one that bridges the cycles offered by Reading and Reid.

THE RESOURCES

The activities and interactive applets were designed to take small steps in developing understanding of a concept, building the ideas through a series of activities. For example, an activity drawing a bead from an unknown population of blue and white beads introduces students to long run relative frequency, where they can visualize the unpredictability of estimates for the proportion of blue beads in the population by sampling a small number of beads and how the estimates become more predictable as the number grows. The technology allows students to make conjectures and investigate patterns as they can quickly and often repeat the sampling from the same and from different populations.

The concept of randomness is revisited in different settings (Oehrtman, 2008). Randomly selecting students from a class to turn in their homework each day makes clear that in small samples, random does not necessarily mean "fair" in terms of outcome but it does mean fair in terms of opportunity (e.g., Kong handed in four assignments over the two-week period, but Charlyne never handed one in). To visualize the representativeness of random samples, students compare random samples of the outcomes of two spinners, one based on the sum of the faces of two fair dice and the other on the outcomes of a spinner with 11 equally likely outcomes. Drawing a sample of maximum speeds from a population of the speeds of different animal types allows students to see again how a sample represents the population from which it was drawn and repeatedly generating different sample sizes makes visible that a small sample size is likely to have a larger variability than a larger sample size. Simulating sampling distributions for a sample statistic such as a proportion allow students to recognize that the variability is relatively constant for repetitions of the same sample size but decreases as the sample size increases (Figure 1).

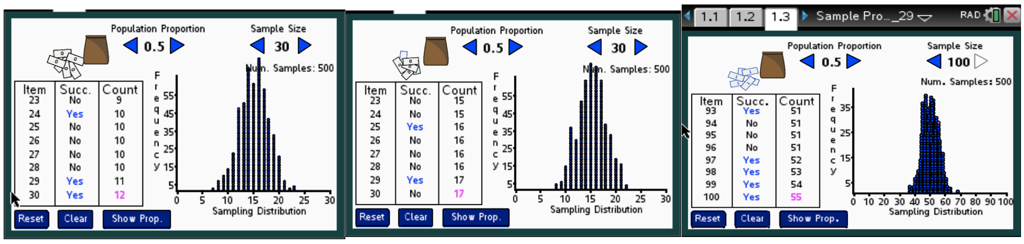

Figure 1. Simulations of sample size 30 and 100 for population that is 50 percent female

The tasks used were designed in light of the research related to student learning, challenges and misconceptions. The misconceptions included the belief that two samples from the same population will be similar, sampling distributions have the same variability for large and small sample sizes, and the sampling distribution for a sample statistic will look like that of the population (Tversky & Kahneman, 1971). 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). Reasoning about distributions is particularly difficult for students when dealing with sampling distributions. Chance, delMas and Garfield (2004) noted that students were not able to reason about sampling distributions until they had a sound understanding of both variability and distribution.

THE RESEARCH

The project was carried out in three iterations of a semester long statistics course for elementary preservice students who were underclassmen in an elementary teacher preparation program at a large midwestern public university. Results indicated the need for changes in the implementation of the course after each year of implementation. The materials were redesigned in light of these findings, with a greater emphasis on hands on activities before engaging with the technology. This paper reports on the third iteration. All thirteen of the students had selected a mathematics emphasis for their certification; their background in statistics was minimal. Students had their own computers, and they used TI© Nspire software to access applets from the Texas Instruments *Building Concepts: Statistics and Probability* files (https://education.ti.com/en/building-concepts) and later in the course used StatKey (www.lock5stat.com/StatKey/). The class met twice a week in 110-minute sessions for a semester. The goals of the course were to enable students to interpret and make sense of data, in particular data related to education, and to give them tools and strategies for their own teaching.

The research question that is addressed in this paper was "What effect does the use of interactive dynamic applets have on student understanding of sampling distributions. In particular, what effect does the technology seem to have on misconceptions related to these concepts?"

THE DATA AND ANALYSIS

The data for the study consist of records of student comments and approaches as they worked through instructional materials related to the applets, reflections of the instructor about student understanding after each class, quizzes, midterm and final exams, students projects and a final survey related to the use of the applets. The data from each source were coded according to concepts identified by the researcher. The data were analyzed by classifying the responses using a hierarchical performance level based on the SOLO taxonomy (Structure of Observed Learning Outcomes, Biggs & Collis, 1982). The analysis with respect to a specific concept was done in three parts: first, identifying elements or features of a concept that could be associated with the levels in the SOLO taxonomy; second, categorizing examples and student responses used during class with respect to the elements in the taxonomy; and third, summarizing the SOLO levels attained by the students with respect to the concept. Table 1 is adapted from Reading and Reid's Interpretation of the SOLO taxonomy for statistical reasoning (2006) to focus on the development and consolidation of a concept image. Table 2 illustrates some features that might be associated with concept images related to random samples, followed by two examples from the midterm test, one of them a classic problem first proposed by Kahneman and Tversky (1972). Table 3 illustrates concept images related to sampling distributions followed by an example from the midterm test. Limitations on the number of pages for this paper prevent displaying any more of the results.

Table 1: SOLO taxonomy and concept images adapted from Reading and Reid (2006)

SOLO taxonomy level	Description of application to concept image
Prestructural (P)	Does not refer to key elements of the concept
Unistructural (U)	Focuses on one key element of the concept
Multistructural (M)	Focuses on more than one key element of the concept
Relational (R)	Links key elements of the concept; relates concepts to other domains

Table 2: Features associated with a concept image related to random samples

SOLO taxonomy level	Concept images for sampling distributions: Randomness
Prestructural (P)	Represents random samples as those that are "accidental" or outcomes that happen without input from an observer
Unistructural (U)	Identifies sampling strategies that will produce a random sample; uses definition of random sample in specifying samples
Multistructural (M)	Connects randomness to bias and equal opportunity for selection; distinguishes between relative frequency and frequency distribution; interprets variability in random samples from the same population
Relational (R)	Connects randomness to short term unpredictability and long run predictability; relates to sampling distributions of a random variable

Example 1: When asked which of four strategies would produce a random sample, all students were correct with respect to three of the strategies with acceptable reasoning related either to bias or describing why everyone would not have an equal chance to be selected. Their reasons fell into category M. The fourth task was, "To represent the population in a room, everyone chooses a number and those with even numbers are the sample." Slightly more than half of the students agreed this would be a random sample because everyone has "a 50% chance of choosing an even number", giving a Category P reason. The high number of incorrect responses might indicate a remaining area of confusion or could be evidence of inexact wording; three of the students qualified their yes by saying that those in the room would have to draw a number from a bag or use another random device to assign the numbers.

Example 2: Half of all newborns are girls, and half are boys. Hospital A records an average of 50 births per day. Hospital B records an average of 10 births a day. On a particular day, which hospital is more likely to record 80% or more of female births?

a. Hospital A (with 50 births a day)
b. Hospital B (with 10 births a day)
c. The two hospitals are equally likely to record such an event.

Please explain your reasoning. (Reaburn, 2008)

Correct responses were given by 70% of the students (as opposed to the 46% correct found by Reaburn) with responses associated with Category R: "Hospital B (with 10 births a day) – Because the sample size is smaller they're more likely to record a higher percentage of girl births. A small sample size allows for more variation in responses and are more likely to have extreme results whereas a larger sample size is more likely to be true to the actual average of female and male births."; "Hospital B is more likely to record 80% or more female births because their sample size is smaller, leading to higher chance that their proportion of female births to male birth will differ from the population proportion."; "Hospital B is more likely to record 80% or more of female births because they will have a larger margin of error because their sampling size is smaller than Hospital A." Thirty eight percent choose c and offered explanations at the P or U levels, e.g., "the two hospitals are equally likely because 80% is a proportion."

Table 3: Features associated with a concept image related to sample sizes

SOLO taxonomy level	Concept images for relating sampling distributions to populations and samples
Prestructural (P)	Does not distinguish between a population, a sample from that population and the distribution of a sample statistic
Unistructural (U)	Identifies/interprets simulated sampling distribution of a sample statistic
Multistructural (M)	Recognizes that sampling distributions of sample means/proportions will be a normal distribution
Relational (R)	Links sample size to variability, visually compares distributions of sample proportions/means with respect to sample size; uses a simulated sampling distribution to consider whether an observed outcome is likely

Example 3: Students were given a graph of the distribution of a uniform population of numbers from 2 to 12 and the graphs of two samples of size 50; one relatively normal distribution, the other approximately uniform. Part a) of the question asked: Which of the two distributions is more likely to be a random sample from the population? Explain your thinking. Part b) displayed Figures 2 and 3 below and asked: Which of the distributions represented in Figures 2 and 3 below could be a simulated distribution of sample means for a sample of size 200 randomly selected from the original population? Explain.

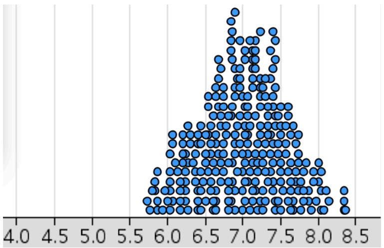

Figure 2

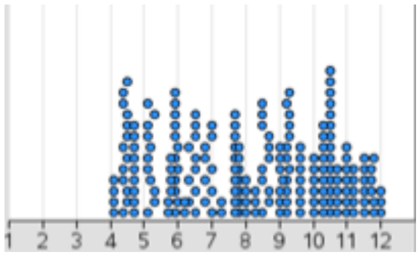

Figure 3

The responses of 92% of the students were correct for part a). Typical explanations were in category M: "This is because when you take random samples from the population above and plot them below it is likely that the distribution will look similar. This figure and the sample are both uniform and if the population was sampled it would likely end up being uniform because points come from anywhere in the graph so the graph will keep a similar distribution." "While it could be either of them because of the small sample size, I'll say it is more likely to be the almost regular sample because it has more of a uniform distribution similar to the population" One response was clearly at the P level confusing random and representativeness: "Neither figure is more likely to be a random sample from the population because if they are random they each have an equally likely chance of occurring."

In responding to part b), the students were divided, with 54% choosing the correct figure (2) and reasoning at the M level, typically the central limit or similar language for their justification. However, 38% argued (one left the answer blank) in ways that indicated they had not transitioned to thinking about the distribution of a sample statistic but were still thinking about the distribution of the sample: "Figure 3 because again all of the number are being selected equally it appears and because for this scenario every number 2 to 12 has an equal chance of being selected." or "I think it could be figure 3 because the x-axis is labeled the same and there would have to be a lot of number taken form [sic] around 6-8 for it to be the figure 2."

RESULTS AND CONCLUSIONS

Other questions from two tests and the final as well as three student projects were analyzed in a manner similar to that described above. In this third implementation, the data suggest that over half of this group of preservice students had at least a multistructural (level M) understanding of random samples and sampling distributions, according to the SOLO taxonomy. The data also suggest areas in which some students struggled, for example, believing that a sampling distribution for a sample statistic should reflect the population distribution.

Much of the literature (Batanero, Burrill, & Reading, 2011) cautions that teachers themselves are often not prepared to teach statistics and may, in fact, leave students with mechanical knowledge but little understanding, misconceptions and doubts. It is important for the field to think through how teacher educators might structure teaching of statistics and probability to lead to increased learning. It is also important to recognize that further research is needed with respect to the use of interactive dynamic technology, which, when focused on developing concepts, might be a possible vehicle to increase understanding and retention of ideas.

REFERENCES

Batanero, C., Burrill, G., & Reading, C. (Eds.). (2011). Teaching statistics in school mathematics - Challenges for teaching and teacher education: *A joint ICMI/IASE Study* (pp. 57–69). New York: Springer.

Biggs, J., & Collis, K. (1982). Evaluating the quality of learning: The SOLO taxonomy. New York, NY: Academic Press.

Chance, B., delMas, R., & Garfield, J. (2004). Reasoning about sampling distributions. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 295–323). Dordrecht, The Netherlands: Kluwer Academic Publishers.

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, 7*(3), www.amstat.org/publications/jse/secure/v7n3/delmas.cfm

Drijvers, P. (2015). Digital technology in mathematics education: Why it works (or doesn't). In: Sung Je Cho (Ed.). *Selected regular lectures from the 12th International Congress on Mathematical Education* (pp. 135–151). Springer, Cham.

Felder, R.M., & Brent, R. (2005). Understanding student differences. *Journal of Engineering Education*, *94* (1), 57–72.

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.

Kahneman, D., & Tversky, A. (1972), Subjective probability: A judgment of representativeness, *Cognitive Psychology, 3*(3), 430–454.

National Research Council. (1999). How people learn: Brain, mind, experience, and school. J. Bransford, A. Brown, & R. Cocking (Eds.). 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. 65–80). http://hub.mspnet.org//index.cfm/19688

Reading, C., & Reid, J. (2006). An emerging hierarchy of reasoning about distribution: From a variation perspective. *Statistics Education Research Journal*, *5*(2), 46–68.

Reaburn, R. (2008). The hospital problem revisited. Tertiary student's perceptions of a problem involving the binomial distribution. In M. Goos, R. Brown, & K. Makar (Eds.) *Proceedings of the 31st Annual Conference of the Mathematics Education Research Group of Australasia*.

Rittle-Johnson, B., Siegler, R., & Alibali, M. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process *Journal of Educational Psychology, 93*(2), 46–36.

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.

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.

Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Bulletin*, *76*(2), 105–110.

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