CONCEPT IMAGES AND STATISTICAL THINKING: THE ROLE OF INTERACTIVE DYNAMIC TECHNOLOGY

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

Understanding a statistical concept involves creating a mental image of that concept. Interacting with carefully developed dynamic interactive technology files can help students build a dynamic "movie clip" of features of the concept that can become the basis for understanding. The paper discusses the impact of interactive "action/consequence" apps used with carefully designed tasks in a project involving elementary preservice teachers. In particular, the technology was used to help students confront misconceptions in making sense of statistical ideas, in particular for the purposes of this paper, variability. Initial results of the analysis of their thinking using a framework adapted from the Solo taxonomy are described.

INTRODUCTION

Coming to understand a mathematical or statistical concept involves creating a mental image of that concept, what Tall and Vinner (1981) call 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 the concept, a student's concept image changes and evolves. For example, a student's first image of standard deviation might be an image of a formula used for computing the standard deviation. If the concept image remains at this level, students will struggle when they are asked to use and interpret standard deviation in different situations. 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).

Interactive dynamic technology can provide students with "live" visualizations of a concept that have the potential to enable students to build robust concept images of the properties, processes and relationships connected to the concept. Applet-like electronic documents can leverage dynamic linkages between representations to address specific learning outcomes for statistical concepts. Interacting with the files allows students to build a concept image that includes dynamic "movie clips" of the features of the concept that can become the basis for understanding. The combination of these documents with inquiry-based activities based on the research about how students come to learn statistical ideas including typical student misunderstandings and misconceptions formed the basis for a statistics course for elementary preservice teachers at a large midwestern university. This paper describes the use of such applet-like documents based on materials from *Building Concepts: Statistics and Probability* (2015) to help students develop robust conceptual structures for key statistical concepts.

BACKGROUND

In 1997, Ben-Zvi and Friedland 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. Initially graphing technology allowed students to visualize mathematical objects and relationships. Dynamic interactive technology when used by a knowledgeable teacher can provide even greater opportunities for students to visualize mathematical and statistical concepts by helping them build robust dynamical conceptual images that can enable them to reason and solve problems from a

In M. A. Sorto, A. White, & L. Guyot (Eds.), *Looking back, looking forward. Proceedings of the Tenth International Conference on Teaching Statistics (ICOTS10, July, 2018), Kyoto, Japan.* Voorburg, The Netherlands: International Statistical Institute. iase-web.org

connected and coherent mathematical structure (Drijvers, 2015). 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). For example, a regression line can be dynamically linked to a visualization of the residual squares and the numerical sum of the squared residuals. 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 (Hodgson, 1996; delMas et al, 1999). 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. Interactive linking supports investigations into varying assumptions and asking "what if" questions that can lead to a better understanding of the concepts involved. Many misconceptions held by students about statistical concepts can be confronted using technology to "predict and-test", 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 will observe (e.g., expected shape of a distribution) and then use the technology to obtain immediate feedback on their understanding.

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. In the applets from *Building Concepts* Oehrtman's three steps are embodied in an "action-consequence-reflection" principle, 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.

THE RESEARCH

The project was carried out in a semester long statistics course for elementary preservice students. The students were underclassmen in an elementary teacher preparation program at a large midwestern public university. Many had selected a mathematics emphasis for their certification; in one class, 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. 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 be literate consumers of statistical data related to education and to give them tools and strategies for their own teaching. A typical class began with a brief overview of key ideas or considerations from the prior class, then a short introduction to a new concept followed by working in small groups through a series of carefully designed questions using the applets to investigate the concept. Students often presented their thinking about the tasks, and discussion focused on making connections across the content they had been learning as well as on considerations that would be important when teaching the concepts to students.

The tasks in the activities were designed in light of the research related to student learning, challenges and misconceptions. They also were constructed following the advice of Black and Wiliam (1998) with respect to formative assessment that tasks can work well only if opportunities for pupils to communicate their understanding are built into the planning. The project focus was to investigate and document how dynamic interactive applets can help learners build robust concept images of key statistical ideas. In particular, research has identified typical misconceptions students have related to core statistical concepts, and the applets and accompanying materials focus on helping students confront these misconceptions by creating visual interactive representations (mini "video clips") that can support a better and more robust understanding of the concept. The research question that is addressed in this paper was **"**What effect does the use of interactive dynamic applets have on typical student misconceptions related to core statistical 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, quzzes, midterm and final exams, students projects and a final survey related to the use of the applets. The data from each source were coded in terms of potential misconceptions. The misconceptions discussed in this paper are related to recognizing and understanding variability in data (Franklin et al, 2006). In particular, students tend to view spread separately from the mean or median, confuse variability with frequency, can not integrate the concepts of variability and shape (Chance, delMas & Garfield, 2004; Groth & Bergner, 2006), lose reference to the context (Cooper & Shore, 2008), and often think of "spread" in data as being evenly distributed (like spreading butter) (Matthews & Clark, 2003; delMas & Liu, 2005). They also believe that sampling distributions for small and large sample sizes have the same variability and that a random sample is a model of the population, thinking only of one observed random sample and not considering the variability across all possible samples and how their sample might fit into that range of possibilities (Chance, delMas, & Garfield 2004).

The applets develop mean as a fair share, found by either taking from the entity with the most and giving to the one with the least until all have the same share or by pooling the objects and dividing the pool equally among the number involved. Measures of variability around the mean are developed by counting the devations from the mean leading to the mean absolute deviation (MAD). Several different applets (adhering to Oertman's third principle of repeated exposure to the ideas in different situations) focus on mean as the balance point for a distribution, where the deviations below the mean "balance" the deviations above the mean. Other applets engage students in estimating the mean and mean +/-MAD for randomly generated distributions of student scores. The progression moves from MAD to standard deviation and from distributions of univariate data with a focus on describing what is typical and what is not to simulated distributions of sample proportions and sample means with a focus on when is a result surprising or not typical.

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 table below is adapted from Reading & Reid's Interpretation of the SOLO taxonomy for statistical reasoning (2006) to focus on the development and consolidation of a concept image.

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

Description of application to concept image
Does not refer to key elements of the concept.
Focuses on one key element of the concept.
Focuses on more than one key element of the concept.
Develops relational links between various key elements of the concept.

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 and linking these to possible misconceptions. 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. The discussion below describes an example of the work related to variability, and Table 2 illustrates some features that might be associated with developing a concept image for variability in data distributions.

Table 2: Features associated with a concept image for variability in data

SOLO taxonomy level	Concept images for variability in data
Prestructural (P)	visualizes variability as total span - range; confuses variability with frequency- images of peaks; visualizes spread as evenly distributed
Unistructural (U)	visualizes variability as spread or differences in the way the data are distributed; connects central clusters to variability; can compute mean absolute deviations, standard deviations, interquartile ranges but cannot associate these visually with a distribution
Multistructural (M)	visualizes variability in terms of deviations from the mean or median; interprets variability in terms of context; recognizes sampling variability- has mental images of different samples of the same size from a population; integrates concepts of variability and shape
Relational (R)	links variability in samples to sample size, in particular visually able to compare sampling distributions of sample proportions or sample means with respect to sample size; connects variability to margin of error with a visual image of interval for plausible populations; correctly interpret margin of error in a context

Example 1:

The following question, typical of ones asked in the literature, presented four histograms and asked: "The plots show the distributions of the number of pairs of shoes owned by four different classes of students. Rank the distributions in order from the one with the least variability to most variability in the number of pairs of shoes per person. Justify your ranking."

Correct ranking and justification would place the task in category M of the taxonomy because students would be expected to interpret variability in terms of context, recognize and describe variability in terms of measures of center and spread, and integrate variability with the shape of a distribution. Responses placed in P typically referred to range ("…the difference between the max and min is very small; "there are many observations covering a wide range.") Those in category U typically mixed correct statements with something not quite correct. ("A has the most variability because the graph is very wide. 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.") The results in Table 3 show that 71% of the students were at level M.

Table 3: Categorization of responses with respect to ranking histograms in terms of variability

SOLO taxonomy level	Ranking histograms in terms of variability with justification
Prestructural (P)	10%
Unistructural (U)	19%
Multistructural (M)	71%
Relational (R)	

Example 2:

The preservice class was given five students' achievement on a state assessment (Figure 1).

Figure 1 Individual student scores on state assessment

The results for individual students are reported to teachers in charts such as the one above (Michigan Department of Education, 2016). The margin of error is represented in the grey bar around the black vertical segment (the segments above each vertical segment highlight these grey bars). The exam question asked for which of the three students A, B, or C, is the margin of error most problematic in terms of the student's level of proficiency? Explain why.

Correct responses to this item were coded R because the answers involved connecting variability to margin of error from a visual image and interpreting margin of error in a context. A typical correct response was "The margin of error is most problematic for student B with a scale score of 2110. His scale score places him within the proficient group. But the margin of error means that he could actually be in the partially proficient group or the advanced group. This is extremely problematic because the student could be placed in a group that he does not actually fit in." Table 4 shows how the results of coding the responses.

Table 4: Categorization of responses with respect to margin of error

SOLO taxonomy level	Ranking histograms in terms of variability with justification
Prestructural (P)	14% (most associated margin of error with lowest score)
Unistructural (U)	10% (misread graphs)
Multistructural (M)	28% (correct reasoning but wrong choice)
Relational (R)	48%

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. The data suggest that over half and tyically up to two thirds of the preservice students had at least a multistructural (level M) understanding of variabilty, according to the SOLO taxonomy. As such, these students did not seem to exhibit the typical misconceptions associated with variability, which is related to the first research question: What effect does the use of interactive dynamic applets have on typical student misconceptions related to core statistical concepts. Approximately one third of the students still struggled with some of the features of variability, often due to language and communication issues, which was evident in their project work.

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 what teacher educators can do to successfully implement teaching of statistics and probability as outlined in the Common Core State Standards and in GAISE (2006). 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.

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.

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

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

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

Chance, B., delMas, R., & Garfield, J. (2004). Reasoning about sampling distribution. 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.

Cooper, L., & Shore, F. (2008). Students' misconceptions in interpreting center and variability of data represented via histograms and stem-and-leaf plots. *Journal of Statistics Education 16*(2). www.amstat.org/publications/jse/v16n2/cooper.html

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

delMas, R., & Liu, Y. (2005, May). Exploring students' conceptions of the standard deviation. *Statistics Education Research Journal*, *4*(1), 55–82. Retrieved July 15, 2007, from http://www.stat.auckland.ac.nz/~iase/serj/SERJ4(1)_delMas_Liu.pdf

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. Springer, Cham

Franklin, C., Kador, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., Scheaffer, R. (2006). Guidelines for assessment and instruction in statistics education (GAISE) report: A Pre-K–12 curriculum framework. Alexandria VA: American Statistical Association.

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.

Hodgson, T. (1996). The effects of hands-on activities on students' understanding of selected statistical concepts. In E. Jakbowski, D. Watkins, & H. Biske (Ed.), *Proceedings of the eighteenth annual meeting of the North American chapter of the international group for the psychology of mathematics education* (pp. 241-246).

Matthews, D., & Clark, J. (2003). Successful students' conceptions of mean, standard deviation and the central limit theorem. Unpublished paper. Retrieved October 20, 2007, from http://www1.hollins.edu/faculty/clarkjm/stats1.pdf.

Michigan Department of Education. (2016). M-STEP Final Reports Webcast.

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. http://hub.mspnet.org//index.cfm/19688

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. Schwier, R.A., & Misanchuk, E.R. (1993). *Interactive multimedia instruction.* Englewood Cliffs, NJ: Educational Technology Publications

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

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*, 151-169.