Cognitive Resources in Student Reasoning about Mean Tendency

Kelly Findley
Florida State University

Jennifer Kaplan
University of Georgia

The ability to conceptualize the sample mean as having a distribution is essential to the development of statistical reasoning. Considerable research on student thinking exists on this topic, but this literature largely assumes a misconception model. This study takes a grounded theory approach to investigate the cognitive resources incoming students possess to reason about sampling distributions and mean tendency. This preliminary report includes data from a pilot study with one student enrolled in an introductory statistics course. She completed both a pre- and post-instruction interview that involved prompts about the distribution of the ages of pennies in circulation and related questions about average ages of groups of pennies. We identify several cognitive resources elicited by the pre- and post-interviews, consider the influence of instruction on the activation of these resources, and briefly discuss implications to statistics teaching. Finally, we outline next steps for data collection with 8-10 students.

Keywords: Statistics, Sampling Distributions, Resources, Constructivism

The heart of statistical reasoning in the introductory statistics course is linking the core concepts of sampling, variability, and distribution to a unified conceptualization of sampling distributions and inference (Garfield, 2002). Unfortunately, superficial knowledge of this and other central ideas results in students resorting to procedural, cookbook approaches to tasks involving statistical inference (Garfield & Ben-Zvi, 2008; Garfield & Zieffler, 2012). Sampling distributions and mean tendency have been identified as extremely difficult concepts (Chance, delMas, & Garfield, 2004; Lunsford, Rowell, & Goodson-Epsy, 2006). Through their research on student thinking involving sampling distributions, Chance and colleagues offered specifics regarding what students should understand about sampling distributions, what they should do with that knowledge, and common misconceptions they may hold. Many pieces in the literature on this topic have aligned with a misconception framing (e.g., Posner, Strike, Hewson, & Gerzog, 1982) and view learning as adoption of expert thinking and replacement of novice thinking (e.g., delMas, Garfield, & Chance, 1999; Garfield, Le, Zieffler, and Ben-Zvi, 2015; Sotos, Vanhoof, Van den Noortgate, and Onghena, 2007).

This study takes a different perspective by seeking to identify cognitive resources students apply to reason about mean tendency. Smith, diSessa, and Roschelle (1993) define *resources* as designating “any feature of the learner's present cognitive state that can serve as significant input to the process of conceptual growth” (p. 124). Resources may be visualized as “knowledge-in-pieces,” representing fine-grained intuitions drawn from experiences and activated in multiple contexts where the learner identifies potential connection (diSessa, 1988). We seek to identify resources students use to reason about mean tendency before formal instruction and how resource activation is influenced by formal instruction. Our research questions are as follows:

1) What cognitive resources do students activate when reasoning about mean tendency?
2) How does formal statistics instruction influence student activation of resources on these topics?

Conceptual Framework

The statistics education literature has traditionally framed novice reasoning about probability and sampling distribution in terms of misconceptions (e.g., ASA, 2016; Cohen, Smith, Chechile, Burns, & Tsai, 1996; Garfield & Ahlgren, 1988; Kahneman, Slovic, & Tversky, 1982; Sotos et al., 2007). *Misconceptions*, as defined by Sotos and colleagues, may represent "any sort of fallacies, misunderstandings, misuses, or misinterpretations of concepts, provided that they result in a documented systematic pattern of error" (p. 99). Misconceptions may be defined to represent any number of ideas, but, a misconception framing has implications on how we view learning, and thus how we define effective teaching. This framing suggests that students have incorrect, stable conceptions of statistical ideas under what Hammer (2004) would term a "unitary" model. In contrast to a unitary model, Hammer suggests we view learners as possessing a "manifold" model that is more fine-grained. Within a manifold model, we will see students applying various resources to arrive at conceptions and conclusions. We view these two alternative perspectives not necessarily as looking at different mountains, but often as looking at the same mountain from different sides. A unitary model tends to look at the outcomes and conclusions the learner makes by synthesizing cognitive resources, while the manifold model searches for starting places and seeks to isolate the learner's cognitive resources.

Some research on student misconceptions suggests that effective instruction involves "confronting" or "eradicating" these misconceptions (e.g., delMas et al., 1999; Eaton, Anderson, & Smith, 1984). For example, delMas and colleagues found it was beneficial to present students with an anomaly to help them understand the Central Limit Theorem (CLT). The researchers modified their simulation program to encourage students to compare predictions with the actual shape of the sampling distribution. While such strategies may have resulted in students accepting the CLT as a fact, it is difficult to ascertain whether students in the study restructured their knowledge cognitively to develop deep and lasting understanding of the CLT.

In his study on student reasoning about probability with dice, Pratt (2000) took a resource perspective. He noted that resources may be contextually appropriate (e.g., the more data we collect, the more stable the results will be) or inappropriate (e.g., the next observation will be 'random' because I cannot steer or control the result). He noted the contextually inappropriate resources were based on short-term, "local" observations that emphasized randomness and unpredictability. The appropriate resources were founded in long-term "global" observations and recognition of probabilistic patterns. Pratt's work contributed to research on children's understanding of probability, but has remained relatively undeveloped in understanding college students' reasoning regarding statistical inference. We intend to work towards that goal.

Methods

Setting

This study is ongoing and taking place at a large public university in the southeastern United States. The target population is students enrolled in introductory applied statistics courses for non-majors. This preliminary report discusses findings from a pilot study with a sophomore Biology major at the university. Karen (pseudonym) was enrolled in a small introductory statistics course for Biology majors and reported having limited high school exposure to statistics content. The instructor of her college course was a Chinese teaching assistant enrolled in the Ph.D. program in statistics, in his third semester of teaching this course.

Data Collection

Data include a 20-minute pre-instruction and 30-minute post-instruction interview with Karen, field notes from the first author's observations of the class periods, and two interviews with the instructor. All interviews were video recorded and transcribed, and hand-written work from Karen's interviews were also kept as data.

In each of the interviews, Karen was provided with an x and y axis template and asked:

> Think about the age of pennies in circulation, like pennies in cash registers or people's money wallets and purses. What is the range of penny years that we would see in circulation? Draw a line to represent how many pennies you would expect there to be across the range of penny years.

Karen was also asked:

> Now think about if we were to take 5 pennies randomly from circulation and find their average age. If we repeatedly did this and collected a list of 5-penny averages, then what would be the range of averages we would see? Draw a line to represent how many of each average would we see.

The third prompt was a variation of the second prompt with 25 pennies instead of 5. The interviewer (first author) asked clarification questions and probed Karen's reasoning during the interviews. Karen answered the same prompts in both interviews.

The instructional period of interest spanned from the introduction to distributions through the end of instruction on the Central Limit Theorem. Field notes from observations focused on the kinds of tasks students completed and the knowledge that was privileged. Interviews were completed with the instructor to capture both reflections about his instruction and his perceived goals for students. These interviews served to triangulate observations and limit bias in the first author's account of the events taking place in class.

Methods of Analysis

We take a grounded-theory approach to identify resources and search for links between formal instruction and student reasoning after instruction. To answer our first research question, we examined the pre-instruction interview for resources by identifying specific statements Karen made and relating them to more general kinds of reasoning (e.g., "I think a larger sample will buffer out the line" was interpreted as "more data means more accuracy"). We tagged spots that were ambiguous, often representing points when Karen's conceptual structure of distributions and mean tendency was potentially underdeveloped or self-contradicting. Using NVivo software, we applied general codes to interview statements and identified three resources.

To answer the second research question, we open-coded Karen's pre-instruction interview, focusing first on the elements of distribution to which Karen attended and the order to which she attended to those things (e.g., the range, the center/middle, the height, the extremes, etc.). We consulted the characteristics from Arnold and Pfannkuch (2015) to guide our organization of these elements. The post-instruction interview was coded in a similar fashion. We compared and contrasted Karen's reasoning before and after instruction to understand how the instruction influenced her reasoning. We also attempted to link elements of the instruction to reasoning approaches and explanations Karen provided in the second interview. We created a causal network diagram to relate Karen's reasoning before and after instruction through elements of the instruction and to synthesize these connections.

Throughout the process, we consulted tactics for generating meaning from Miles, Huberman, and Saldaña (2014). For example, we began the analysis by coding to look for patterns and themes, we clustered and partitioned codes, made comparisons and contrasts across Karen's pre and post interview, subsumed particular statements from the interviews into potential resources, and worked to make conceptual coherence of Karen's reasoning.

Findings

Figure 1 shows Karen's drawings for the population distribution (on the left) and the sampling distribution for n=5 (on the right).

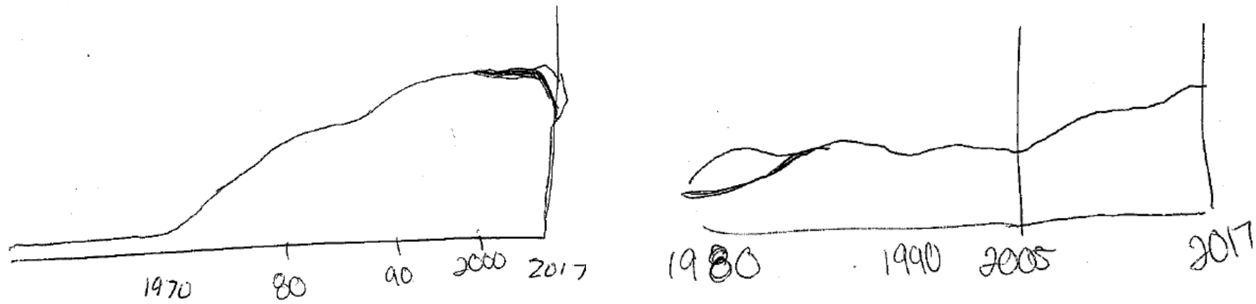

Figure 1. Karen's pre-interview drawings for the population and for sampling distribution with n=5.

Resources

Average is middle. This was the first resource Karen used as she reasoned about the second prompt. She fabricated a likely sample she could imagine pulling from her change purse and thought about the average of that sample: "Let's say you get 5 that are like 1980, 1987, 1992, 1997, and then 2005, then it would be like, the average between that, so, like somewhere around 1993." This resource served as a starting point for reasoning about individual sample means, but was no longer activated when reasoning about the shape of the sampling distribution.

The sample will resemble the population. Karen said the following while drawing the sampling distribution for n=5: "Shapewise, [the population and sampling distribution are] not too different, like I have the lower down here and then it increases over time, and then it's kind of like a dip." In this description, Karen recognizes a relationship between these two distributions. What she is noting is a similarity in shape, with the sampling distribution (in her view) being a smoother, slightly flatter version of the population.

Larger sample will produce more consistency. This resource was activated as she continued to reason about the second prompt, and again as she reasoned about the third prompt with n=25. "I feel like this [the population distribution] would be less…like consistent than this [sampling distribution for n=5] would be, I feel like this [sampling distribution for n=5] would be not straight but straighter of a line." When reasoning about the distribution of sample means from even larger samples, she said, "if you take 1000 pennies, if there's not that many outliers, or if it's like evenly, dispersed number of years I guess, it would kind of make it a flatter line." She also mentioned penny averages being "more likely to change" when coming from small samples, thus viewing sample size as a sort of weighting component to variability in the mean.

Two Resource Perspectives

In trying to make sense of the path of reasoning that Karen took, we identified two perspectives of reasoning. The first perspective focuses on the individual pennies and the influence of the population on the samples. This perspective was used frequently in early statements as Karen reasoned about individual pennies that could plausibly be in a sample she took from her change purse. The resources associated with this perspective were "the sample will resemble the population" and "average is middle."

The second perspective focuses more on abstract patterns. When Karen was asked to reason about larger samples, she shifted to thinking about "larger sample means more consistency." As a result, Karen believed there was a correct answer that the sampling distribution shape was approaching. Because she could no longer reason about specific plausible samples, she resorted to this more abstract resource.

Instructional Influence on Karen's Activation of Resources

Consistent with a constructivist perspective on learning, we view Karen's pre-instruction reasoning as representing an existing conceptualization of the relevant statistics content. Even though this knowledge structure was inconsistent and flexible, we do not believe Karen's conceptualization was a "blank slate" ready to receive and adopt correct knowledge structures. We view the formal instruction as a filter on Karen's reasoning and resource activation. Such a framing is neutral: the filter can be beneficial by refining, challenging, or introducing new resources, but this filter can also be detrimental by discouraging resource activation and promoting rote memorization with no deeper conceptual connection.

As she began to reason about the sampling distribution in the post-instruction interview, Karen frequently activated the resource that "average is middle." She pulled heavily on the Central Limit Theorem as evidence for why averages cluster in the middle, but as a definition absent of cognitive conviction. When probed to compare her initial interview drawings to her current ones, Karen recalled the previous resource of "larger samples produce more consistency" and her belief that larger samples will "buffer out" the sampling distribution to a flat line. At this point, Karen struggled to reconcile these seemingly contradicting resources. At the close of the interview, she questioned whether there might be a distinction between a frequency distribution and a probability distribution. Overall, Karen had not radically changed her thinking; she instead appeared to be suppressing a key resource, "larger samples produce more consistency," because it did not align with what she learned in class. "Average is middle," however, could still be aligned with the instruction.

Implications and Future Work

On the surface, it is easy to miss that Karen did not cognitively adopt the Central Limit Theorem. Instead, it was a fact that she could articulate. While she attempted to justify it with one resource she had, she suppressed other resources with the potential to do so. If this pattern can be generalized, it is possible Karen might no longer attempt to reconcile future statistical content with her other resources when such a conflict exists and, instead, resort to a "cookbook" approach to statistical inference. Therefore, arming instructors with a list of relevant resources rather than a list of common misconceptions might lead to more cognitively robust reasoning and avoid leading students to a "cookbook" approach to statistical inference.

In the fall, we will conduct interviews with 8-10 students as we attempt to test and refine the resources we identified, search for others, look for negative cases, and make if-then tests about student reasoning in the instruction. Students will be pooled from a large-lecture introductory course, and students will again be interviewed before and after relevant instruction.

Audience Questions

- Are there papers we are not aware of in the statistics education literature that take a resource framing of learning?
- In looking at our evidence, do you agree with the resources that we have identified?
- Are there benefits to still identifying misconceptions in student thinking on this topic?

References

American Statistical Association. (2016). *Guidelines for Assessment and Instruction in Statistics Education: College Report* 2016. http://www.amstat.org/education/gaise.

Arnold, P., & Pfannkuch, M. (2015). Describing distributions. In K. Makar, B. de Sousa, & R. Gould (Eds.), *Sustainability in statistics education. Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS9, July, 2014), Flagstaff, Arizona, USA.* Voorburg, The Netherlands: International Statistical Institute.

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

Cohen, S., Smith, G., Chechile, R. A., Burns, G., & Tsai, F. (1996). Identifying impediments to learning probability and statistics from an assessment of instructional software. *Journal of Educational and Behavioral Statistics*, *21*(1), 35-54.

delMas, B., 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).

diSessa, A. A. (1988). Knowledge in pieces. In Forman, G. & Pufall, P. B. (Eds.), *Constructivism in the computer age* (pp. 49-70). Hillsdale, NJ: Erlbaum.

Eaton, J. F., Anderson, C. W., & Smith, E. L. (1984). Students' misconceptions interfere with science learning: Case studies of fifth-grade students. *The Elementary School Journal*, *84*(4), 365-379.

Garfield, J. (2002). The challenge of developing statistical reasoning. *Journal of Statistics Education*, *10*(3).

Garfield, J., & Ahlgren, A. (1988). Difficulties in learning basic concepts in probability and statistics: Implications for research. *Journal for Research in Mathematics Education, 19*(1), 44-63.

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

Garfield, J., Le, L., Zieffler, A., & Ben-Zvi, D. (2015). Developing students' reasoning about samples and sampling variability as a path to expert statistical thinking. *Educational Studies in Mathematics*, *88*(3), 327-342.

Garfield, J., & Zieffler, A. (2012). Developing statistical modelers and thinkers in an introductory, tertiary-level statistics course. *ZDM*, *44*(7), 883-898.

Hammer, D. (2004). The variability of student reasoning, lecture 3: Manifold cognitive resources. In E. Redish & M. Vicentini (Eds.), *Proceedings of the Enrico Fermi Summer School, Course CLVI* (pp. 321-340). Bologna, Italy: Italian Physical Society.

Kahneman, D., Slovic, P. and Tversky, A. (1982), *Judgement under uncertainty: Heuristics and biases*, Cambridge, UK: Cambridge University Press.

Lunsford, M., Rowell, G. H., & Goodson-Espy, T. (2006). Classroom research: Assessment of student understanding of sampling distributions of means and the central limit theorem in post-calculus probability and statistics classes. *Journal of Statistics Education*, *14*(3).

Miles, M., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods* sourcebook. Thousand Oaks, CA: SAGE.

Pratt, D. (2000). Making sense of the total of two dice. *Journal for Research in Mathematics Education*, *31*(5), 602-625.

Smith III, J. P., diSessa, A. A., & Roschelle, J. (1994). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *The Journal of the Learning Sciences, 3*(2), 115-163.

Sotos, A., Vanhoof, S., Van den Noortgate, W., & Onghena, P. (2007). Students' misconceptions of statistical inference: A review of the empirical evidence from research on statistics education. *Educational Research Review, 2*, 98–113.