Conceptual issues in understanding sampling distributions and margins of error[1]

(Synopsis of Phase 1 of NSF Project *Multiplciative Reasoning as a Foundation for Stochastic Reasoning)*

Patrick W. Thompson
Luis Saldanha

Vanderbilt University

Statistics and probability have been relatively minor topics in school mathematics, both at elementary and secondary levels. Of 39557 content blocks analyzed in the TIMSS curriculum study of US 4th- and 8th-grade math texts, only 6% involved data representation or data analysis as a primary idea. Even then, 90% of those instances occurred in exercises or in diagrams associated with unrelated topics; only 10% of them (0.6% of the total) involved narrative associated with the central ideas. With probability the situation was more extreme. Two percent of the content blocks involved uncertainty or probability as a primary idea, and of those only 9% (0.2% of the total) occurred in related narrative.[2] Moreover, when they were the explicit topic of study the focus was calculating mean, median, and mode for sets of numbers, making bar charts from data sets, or calculating "number of successful outcomes divided by number of possible outcomes".

In like fashion, the preponderance of studies through the early 1990's on statistical and probabilistic reasoning, though relatively small in number, focused on students' understanding of center of a data set or determining numbers that somehow represented a data set (Mokros & Russell, 1995; Shaughnessy, 1992; Strauss & Bichler, 1988) and on representing and interpreting statistical graphs (Wainer, 1992). Connections with inferential statistics were examined only indirectly (Konold, Pollatsek, Well, Lohmeier, & Lipson, 1993), such as examining peoples' intuitive reasoning strategies related to population proportions and judgements one might make about them (Nisbett, Krantz, Jepson, & Kunda, 1983). Research on learning and teaching statistics as data analysis or its connections with probabilistic reasoning had not focused on conceptual operations and imagery ("ways of thinking and understanding") that might support coherent, sound stochastic reasoning.

At the same time that issues related to statistics or probability hardly appeared in textbooks or state curriculum guides, professional and business organizations began calling for dramatically increased attention to students' statistical reasoning (NCTM, 1989, 1998), as did an increasing number of mathematics and statistics educators (Moore, 1991; Shaughnessy, 1992; Shaughnessy, Garfield, & Greer, 1996). Recently, research on stochastic learning and reasoning has had greater theoretical orientations, such as the use of Vygotskian frameworks to take into account broader contexts of learning and teaching (Gordon, 1995), but we are still in the infancy of understanding issues of stochastics pedagogy and reasoning as a coherent conceptual field.

Recent research has begun to address the question of what ideas and ways of thinking might be central to students' developing mature stochastic reasoning. For example, Watson and Moritz (2000) used cross-sectional data to investigate students' developing understandings of sample and sampling. Shaughnessy and colleagues (Shaughnessy, Watson, Moritz, & Reading, 1999) investigated conditions under which students would interpret situations with the understanding that outcomes will probably be different if the same situation were examined with a different set of data.

This study investigated students' abilities and difficulties to conceive the ideas of a statistic's sampling distribution and of margin of error. Twenty-seven 11^{th}- and 12^{th}-grade students, enrolled in a non-AP semester-long statistics course, participated in a nine-session teaching experiment addressing ideas of sample, sampling distributions, and margins of error. Our aim was to produce conceptual analyses of these ideas (Glasersfeld, 1995; Steffe, 1996; Thompson, 2000) – ways of thinking about them that are schematic, imagistic, and dynamic – and hypotheses about their development in relation to students' classroom engagement.

Three research team members were in the classroom during all lessons: the first author designed and conducted the instruction; the second author observed the instructional sessions and took field notes; a third member operated the video cameras. Although an a priori outline of the intended teaching/learning trajectory guided the progress of the experiment, instructional activities and lessons were revised according to what the research team perceived as important unforeseen issues that arose for students in each instructional session. Such revisions occurred on a daily basis as a result of the team's meeting immediately after each session to discuss what had transpired in that session and what possible directions to take for the next session(s).

The teaching experiment began with a focus on news reports that mentioned data about sampled populations and news reports about populations per se (raising the issue of sampling variability). The experiment then progressed to questions of "what fraction of the time would you expect results like these?" This entailed having students employ, describe the operation of, and explain the results of computer simulations of taking large numbers of samples from various populations with known parameters. The experiment ended by examining simulation results systematically, with the aim that students see that sampling distributions are largely unaffected by underlying population parameters, but are affected in important ways by sample size.

Students' understandings were investigated in three ways: (1) Tracing their participation in classroom discussions (all instruction was videotaped); (2) Written, post-experiment examination; (3) Post-experiment individual interviews.

Analyses revealed that better performing students and students exhibiting coherent discourse during class had developed a multi-tiered scheme of conceptual operations centered around the image of repeatedly sampling from a population, recording a statistic, and tracking the accumulation of statistics as they distribute themselves along a range of possibilities. These operations, in turn, seemed to be grounded in an image of samples as quasi-proportional, mini-versions of the sampled population[3]. Poorer-performing students differed from this in significant ways. (1) They tended to view samples additively, simply as some of the population. (2) Their sense of variability did not extend to ideas of distribution. Instead, they understood that sample statistics vary, but only to the extent that if we were to draw some more samples and compute statistics from them, those statistics would be different from the ones already drawn. (3) Their ability to coordinate various levels of activity (drawing one sample and calculating a statistic, repeating this process many times, analyzing outcomes from the second-level process, etc.) was significantly weaker than was the higher-performing students'.

These results are significant in a number of ways. First, they have direct implication for the design of curriculum and instruction in high school and college statistics courses. Second, they point to issues that can be addressed in lower grades (e.g., proportionality and multi-level situations) that might remove obstacles that are encountered by students not so prepared. Other implications will be explored in the full paper and in the presentation.

References

Glasersfeld, E. v. (1995). Radical constructivism: A way of knowing and learning. London: Falmer Press.

Gordon, S. (1995). A theoretical approach to understanding learners of statistics. Journal of Statistics Education, 3(3)(http://www.amstat.org/publications/jse/v3n3/gordon.html).

Konold, C., Pollatsek, A., Well, A., Lohmeier, J., & Lipson, A. (1993). Inconsistencies in students' reasoning about probability. Journal for Research in Mathematics Education, 24(5), 392–414.

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

Moore, D. S. (1991). Statistics: Concepts and controversies. (3rd ed.). New York: W. H. Freeman.

NCTM. (1989). Curriculum and evaluation standards for school mathematics. Reston, VA: National Council of Teachers of Mathematics.

NCTM. (1998). Principles and standards for school mathematics (Discussion draft). Reston, VA: National Council of Teachers of Mathematics.

Nisbett, R. E., Krantz, D. H., Jepson, C., & Kunda, Z. (1983). The use of statistical heuristics in everyday inductive reasoning. Psychological Review, 90, 339–363.

Shaughnessy, J. M. (1992). Research in probability and statistics: Reflections and directions. In D. Grouws (Ed.), Handbook for research on mathematics teaching and learning (pp. 465–494). New York: Macmillan.

Shaughnessy, J. M., Garfield, J., & Greer, B. (1996). Data handling. In A. J. Bishop (Ed.), International handbook of mathematics education (pp. 205-237). Dordrectht, The Netherlands: Kluwer.

Shaughnessy, J. M., Watson, J., Moritz, J., & Reading, C. (1999, April). School mathematics students' acknowledgement of statistical variation. Paper presented at the Research presession of the annual National Council of Teachers of Mathematics meeting, San Francisco.

Steffe, L. P. (1996). Radical constructivism: A way of knowing and learning [Review of the same title, by Ernst von Glasersfeld]. Zentralblatt für Didaktik der Mathematik [International reviews on Mathematical Education], 96(6), 202-204.

Strauss, S., & Bichler, E. (1988). The development of children's concepts of the arithmetic average. Journal for Research in Mathematics Education, 19(1), 64-80.

Thompson, P. W. (2000). Radical constructivism: Reflections and directions. In L. P. Steffe & P. W. Thompson (Eds.), Radical constructivism in action: Building on the pioneering work of Ernst von Glasersfeld (pp. 412-448). London: Falmer Press.

Wainer, H. (1992). Understanding graphs and tables. Educational Researcher, 21(1), 14–23.

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

[1] Research reported in this paper was supported by National Science Foundation Grant No. REC-9811879. Any conclusions or recommendations stated here are those of the author and do not necessarily reflect official positions of NSF.

[2] These results are derived from the TIMSS U. S. curriculum data, available at http://ustimss.msu.edu/data/individual/dat840.zip.

[3] This is not to say that they *anticipated* such a relationship actually existing between any given sample and the population from which it came. Rather, they anticipated that *many* samples would resemble the population "more or less".