Liu, Y., & Thompson, P. W. (2004). Teachers' personal and pedagogical understanding of probability and statistical inference. In D. McDougal (Ed.), *Proceedings of the Twenty-sixth Annual Meeting of the International Group for the Psychology of Mathematics Education*. Toronto: PME-NA.

TEACHERS' PERSAONAL AND PEDAGOGICAL UNDERSTANDING OF PROBABILITY AND STATISTICAL INFERENCE

Yan Liu
Vanderbilt University
Yan.liu@vanderbilt.edu

Pat Thompson
Vanderbilt University
Pat.Thompson@vanderbilt.edu

The purpose of this study is to develop an insight into teachers' personal and pedagogical understanding of probability and statistical inference. To this end, we undertook a teaching experiment with eight high school mathematics teachers. The teaching experiment was designed with the purpose of provoking the teachers to express and to reflect upon their instructional goals, objectives, and practices in teaching probability and statistics. Our results indicated that teachers had a complicated, inconsistent mix of meanings with regard to the ideas of probability, sampling distribution, hypothesis testing, and margin of error.

STATEMENT OF PROBLEM

Teachers' understanding of significant mathematical ideas has profound influence on their capacity to teach mathematics effectively (Ball & McDiarmid, 1990; Ball, 1990; Ball & Bass, 2000; Borko et al., 1992; Eisenhart et al., 1993; Simon, 1994; Sowder, Philipp, Armstrong, & Schapelle, 1998; Thompson, 1984; Thompson & Thompson, 1996), and, in turn, on what students end up learning and how well they learn (Begle, 1972; 1979). This has important implications for how teacher educators think about ways of supporting teachers' professional development—that supporting transformation of teaching practices takes careful analysis of teachers' personal and pedagogical understandings of what they teach.

Probability and statistical inference are among the most important and challenging ideas that we expect students to understand in high school. They have had an enormous impact on scientific and cultural development since its origin in the mid-seventeen century. The range of their applications spread from gambling problems to jurisprudence, data analysis, inductive inference, and insurance in eighteen century, to sociology, physics, biology and psychology in nineteenth, and on to agronomy, polling, medical testing, baseball and innumerable other practical matters in twentieth (Gigerenzer et al., 1989). Along with this expansion of applications as well as the concurrent modification of the theories themselves, probability and statistical inference have shaped modern science and transformed our ideas of nature, mind, and society. Given the extraordinary range and significance of these transformations and their influence on the structure of knowledge and power, and on issues of opportunity and equity in our society, the question of how to support the development of coherent understandings of probability and statistical inference takes on increased importance.

There have been many investigations of ways people understand probability and statistical inference. Psychological and instructional studies consistently documented poor understanding or misconceptions of these ideas among different populations across different settings (Fischbein & Schnarch, 1997; Kahneman & Tversky, 1973; Konold, 1989; Konold, Pollatsek, Well, Lohmeier, & Lipson, 1993a). The challenges for students are not associated merely with acquiring new skills but with overcoming ways of thinking that are unpropitious for reasoning probabilistically (Kahneman, Slovic, & Tversky, 1982; Konold, 1989, 1991) or for understanding probability as a mathematical model (Shaughnessy, 1992; Piaget & Inhelder, 1975), and dealing with the pervasive traditional teaching that works against efforts to make sense of stochastic concepts and ideas (Fischbein, 1975). Contrary to the overwhelming evidences of people's difficulties in learning probability and statistical inference, there is a general lack of insight into mechanisms by which transmission of this knowledge in classroom happens. Particularly, research on statistics education has not attended to teachers' understanding of probability and statistics (Garfield & Ben-Zvi, 2003).

The goal of this study is to explore teachers' personal and pedagogical understanding of probability and statistical inference. To this end, we undertook a teaching experiment with eight high school mathematics teachers. This teaching experiment is an early, highly exploratory step of a larger research program which aims to understand ways of supporting teachers learning and their transformations of teaching practices into one that is propitious for students learning in the context of probability and statistics instruction. We designed the teaching experiment with the purpose of provoking the teachers to express and to reflect upon their instructional goals, objectives, and practices in teaching probability and statistics. Our primary goal was to gain an insight into the issues, both conceptual and pedagogical, that teachers grapple with in order to teach probability and statistics effectively in the classroom.

RESEARCH DESIGN

This study is the last teaching experiment in a research project that entailed a total of 5 teaching experiments conducted over a 40-month period and involved three different groups of participants. The prior four teaching experiments investigated high school students' thinking as they participated in classroom instruction designed to support their learning of sampling, probability, and statistical inference as a scheme of interrelated ideas. The aim was to develop epistemological analyses of these ideas (Thompson & Saldanha, 2000)—ways of thinking about them that are schematic, imagistic, and dynamic—and hypotheses about their development in relation to students' engagement in classroom instruction. Using the products and insights we obtained from these previous teaching experiments (Saldanha, 2003; Saldanha & Thompson, 2002; Thompson & Liu, 2002), in this study we engaged a group of high school teachers in rethinking what they hope students learn from statistics instruction and in reflecting on ways of affecting students' learning.

METHODOLOGY

We conducted the study using a combination of design experiment and constructivist teaching experiment methodologies. Design experiment methodology (Gravemeijer, 1994) provides an emphasis on cycles of design, implementation, evaluation, and redesign. Constructivist teaching experiment methodology (Steffe & Thompson, 2000) provides an emphasis on conceptual analysis of mathematical ideas and analyzing subjects' participation in classroom discussions and interviews for cues about their ways of knowing mathematical ideas.

DESIGN AND IMPLEMENTATION

Eight high school mathematics teachers—six female and two male teachers— participated in the workshop. Among these eight teachers, two had 2-3 years of teaching experience, three had 7-9 years of teaching experience, and three had 21-28 years. Three teachers taught AP statistics. Five taught a probability and statistics chapter in high school mathematics course. The research team consisted of the PI (the second author), a collaborating teacher, and three graduate students. The team designed the workshop activities and artifacts during 9 months prior to the workshop. The collaborating teacher hosted most of the workshop activities and conversations. The PI served as an observer of the workshop and occasionally hosted the workshop or participated in the conversation. One graduate student took field notes and managed miscellaneous logistic work. The first author and another graduate student recorded the workshop sessions with front and back cameras, and made observations and notes during the workshop.

The workshop discussion progressed over eight sessions in two weeks. The workshop began at 9am each day and concluded at 3pm, with a 30- minute lunch break. At the end of each day, the research team met briefly to discuss our observations and suggestions on modification of next days' activities. We also made photocopies of teachers' notes at the end of each day. Each teacher was interviewed 3 times for about 45 to 60 minutes each time. Interviews were conducted once before the workshop and at the end of each week. We video recorded all interviews and kept record of teachers' work during the interviews.

We engineered the discussions so that teachers first worked on and discussed the problems as first-order participants. We used these occasions to construct models of teachers' personal understanding of the ideas of probability and statistical inference. We then initiated pedagogical conversations about these ideas—given these ways of understanding these ideas, what are the implications for teaching them? We intended to elicit reflective conversations in the sense that what was previously discussed became objects of thoughts and conversation (Cobb, Boufi, McClain, & Whitenack, 1997). The interviews were designed to include general questions concerning teachers' stochastic reasoning, as well as specific questions that were tailored to each teacher according to our observation and conjectures about his or her knowledge and beliefs.

DATA ANALYSIS

The data for analysis include video recordings of all workshop sessions made with two cameras (36.5 hours) and individual interviews (approximately 24 hours), the teachers' written work, field notes, and documents made during the planning of the workshop. The analytical approach we employed in generating descriptions and explanations was consistent with Cobb and Whitenack's (1996) method for conducting longitudinal analyses of qualitative data and Glaser and Strauss' (1967) grounded theory, both of which highlights an iterative process of generating and modifying hypotheses in light of the data. Analyses generated by iterating this process were aimed to develop increasingly stable and viable hypotheses and models of teachers' understanding.

We began by first reviewing the entire collection of videotaped workshop sessions and interviews. Our primary goal in this process was to develop an overall sense of what had transpired in the workshop, and to identify video segments that seemed potentially useful for gaining insight into one or more teachers' personal and pedagogical understandings of probability and statistical inference. Segments containing direct evidence of teachers' thinking, miscommunication in discussions of problems or ideas, or controversy about mathematical meanings or pedagogical practices were especially significant. Later, these video segments were transcribed. We then annotated the transcripts with the purpose of developing hypotheses of teacher's understanding. The guiding question for making sense of a teacher' utterance is: *What might he have been thinking (or seeing the situation, or interpreting the previous conversation) so that what he said made sense to himself?*

RESULTS

Our analysis indicated that, collectively and individually, teachers had a complicated, inconsistent mix of meanings with regard to ideas of probability and statistical inference.

PROBABILITY

Teachers' understanding of probability covered a broad spectrum:

1) thinking that probability is a (subjective) judgment based on personal experiences;
2) thinking that probability is about predicting the state of a specific completed (or to-be-completed) event about which one does not know the actual result;
3) thinking that probability is about selecting one outcome from a set of possible outcomes;
4) thinking that probability is about imagining a collective of results all generated by a single process that yields results that are more dense in some regions of possible values than in other regions, i.e., a stochastic conception of probability (Thompson & Liu, 2002).

We found that the representation of probability questions or statements has a direct association with teachers' interpretation of probability. When a probability question or statement explicitly states a collection of people or events, e.g., "A study of *over 88,000 women* found that total folate intake was not associated with the overall risk of breast cancer," or "*drivers* with three or more speeding tickets are twice as likely to be in a fatal accident as are drivers with fewer than three tickets," teachers were more likely to interpret probability as a group characteristic. On the other hand, when a probability question or statement is expressed as about a single event, e.g., "*your* risk of being killed on an amusement park ride? One in 250 million." or "what is the probability that *my* car is red?" teachers were less likely to have a background image of a collection of similar events against which the probability is evaluated.
Even when some teachers later did conceive situations stochastically, they did not understand, or were reluctant to accept, that probability is determined by the stochastic processes one imagines/constructs. Teachers argued whether a probabilistic situation could have multiple interpretations and thus have different values depending how one interprets it. The group that argued against it seemed to believe that a probabilistic situation should not subject to multiple interpretations. They did not, or refused to, see a distinction between a situation as it is stated/written and its (possibly multiple and conflicting) interpretation. However, both groups (one who accepts multiple interpretations and one who don't) shared a commitment to "finding one correct answer". They believed that computer simulation could decide a correct answer, without realizing that simulations are designed according to one's interpretation of the underlying stochastic process, and thus only confirms one's interpretation of a probabilistic situation.

SAMPLING DISTRIBUTION

Our previous teaching experiments revealed students' profound difficulties in conceiving a distribution of sample statistics (Saldanha, 2003). Unlike students, teachers seemed to have a good grip on the hierarchical process that generates the distribution of sample statistics. We observed only one instance in which one teacher referred to a proportion of samples as a sample proportion.

Teachers also understood the relationship between sample size and variability, i.e., as the sample size increases, the variability of the distribution of sample statistics decreases. However, it is not a general idea for them in the sense that their understanding was very contextualized (Thompson, Liu, & Saldanha, 2004). For example, teachers examined the distribution of sample proportions of samples of size 100 from a population a proportion of which have a certain characteristic. They repeated the simulation several times taking 150 samples to see how stable the distribution was from trial to trial, then repeated it several times taking 500 samples, then repeated it several times taking 1000 samples. Only one teacher recognized that "taking *x* samples of size 100" was a stochastic process, and therefore that the relationship between sample size and variability that they had just stated also applied to their investigation.

Teacher also had an intuitive understanding that if the population size increases, then the sample size has to increase proportionally in order to maintain the same variability of the distribution.

HYPOTHESIS TESTING

Teachers were unfamiliar with the logic of hypothesis testing. For example, we gave the teachers this scenario: "Assume that sampling procedures are acceptable and that a sample is collected having 60% favoring Pepsi. Argue for or against this conclusion: This sample suggests that there are more people in the sampled population who prefer Pepsi than prefer Coca Cola." (With a list of 135 simulated samples of size 100 from a population split 50-50 in preference). Teachers initially took the position that there was *no* basis for arguing this conclusion because the data confirmed that the population percent was 50% (the null hypothesis). Two Teachers eventually concurred with the workshop leader that the data suggested that samples of 60% or more were sufficiently rare so as to reject the null hypothesis. Some teachers were reluctant to accept this logic. One teacher argued that one could not make any judgment based on the fact that a rare sample occurred, because the sample *could* occur however rare its chance of occurrence might be. This argument revealed the teacher's commitment to having a *single* valid judgment, which is inconsistent with the idea that decision rule does not decide a single valid judgment, but only ensure that the error rate is small over the long run.

In one scenario, when asking to create a decision rule for testing hypothesis, some teachers proposed one hypothetical simulation result that would confirm alternative hypothesis instead of a general principle for rejecting null hypothesis. As a result, they could not reconcile the difference between the hypothetical simulation result and the actual simulation result. Only one teacher realized that hypothesis testing was like proof by contradiction.

MARGIN OF ERROR

We also found a common misunderstanding of margin of error—all the teachers initially believed that a poll result of "76% with margin of error $\pm$4%" (confidence level 95%) meant that 95% of the sample proportions would fall within the interval [76%-4%, 76%+4%]. Most teachers came to see the error in this thinking. During the second interview, 7 out of 8 teachers correctly interpreted margin of error, i.e., 95% of the sample statistics would be expected to fall within 4% point of the true parameter.

After teachers understood the concept of margin of error, they raised an interesting question: "why bother taking a poll if we don't know *for sure* if the poll result will be among those 95% of the poll results that are within a certain interval of the true population parameter?" The conversation around this question revealed that 1) teachers were uncomfortable with the uncertainty that is the very reason for resorting to statistical information, and 2) teachers had a tendency to fixate on individual cases (in this case, the poll result) as opposed to statistical patterns in the long run.

PADEGOGICAL UNDERSTANDING

Teachers exhibited a strong commitment to "finding the right answer". Many of the tasks we designed were conceptually challenging—the answers were not obvious, and sometimes, they may have different answers depending on different interpretations. We found many instances in which teachers could not move on to pedagogical conversation if an agreement on *one* correct answer was not achieved.

Teachers believed that students learn an idea by "hearing it said correctly". When trying to learn ideas from each other during the workshop, teachers seemed to place low confidence in their own understanding, but readily accept ideas uttered by those whom they perceived as experts. For example, when one teacher questioned the validity of "equal likelihood", the other defended it by saying "the formula of probability is based on equal likelihood, therefore it must be valid." This conception of learning projected potential ways of teaching that are insensitive to students' understanding.

This way of teaching was also projected by the ways in which teachers negotiate meanings. We found that teachers often dismissed the opposing point of view as "mistakes", convincing others that they are wrong by telling them they are wrong, instead of trying to understand where they are coming from.

SIGNIFICANCE

This study will make three contributions to our understanding of the teaching and learning of probability and statistics. It provides greater insight into stochastic reasoning as a foundation for understanding probability and statistical inference. It suggests limitations to which we should be alert in high school teachers' understanding. And last, it suggests areas in which teachers' development of personal and pedagogical understanding in probability and statistics needs support.

References

Ball, D., & McDiarmid, G. W. (1990). The subject matter preparation of teachers. In W. R. Houston (Ed.), *Handbook of research on teacher education* (pp. 437-449). New York: Macmillam.

Ball, D. L. (1990). The mathematical understandings that prospective teachers bring to teacher education. *Elementary School Journal, 90*, 449–466.

Ball, D. L., & Bass, H. (2000). Interweaving content and pedagogy in teaching and learning to teach: Knowing and using mathematics. In J. Boaler (Ed.), *Multiple perspectives on mathematics teaching and learning* (pp. 83-106). Stamford, CT: Ablex.

Begle, E. G. (1972). *Teacher knowledge and pupil achievement in algebra* (NLSMA Technical Report No. Number 9). Palo Alto, CA: Stanford University, School Mathematics Study Group.

Begle, E. G. (1979). *Critical variables in mathematics education: Findings from a survey of the empirical literature*. Reston, VA: National Council of Teachers of Mathematics.

Borko, H., Eisenhart, M., Brown, C. A., Underhill, R. G., Jones, D., & Agard, P. C. (1992). Learning to teach hard mathematics: Do novice teachers and their instructors give up too easily? *Journal for Research in Mathematics Education, 23*(3), 194–222.

Cobb, P., Boufi, A., McClain, K., & Whitenack, J. (1997). Reflective discourse and collective reflection. *Journal for Research in Mathematics Education, 28*(3), 258-277.

Cobb, P., & Whitenack, J. (1996). A method for conducting longitudinal analyses of classroom videorecordings and transcripts. *Educational Studies in Mathematics, 30*(3), 213-228.

Eisenhart, M., Borko, H., Underhill, R., Brown, C., Jones, D., & Agard, P. C. (1993). Conceptual knowledge falls through the cracks: Complexities of learning to teach mathematics. *Journal for Research in Mathematics Education, 24*(1), 8–40.

Fischbein, E., & Schnarch, D. (1997). The evolution with age of probabilistic, intuitively based misconceptions. *Journal for Research in Mathematics Education, 28*(1), 96-105.

Garfield, J., & Ben-Zvi, D. (2003). Research on statistical literacy, reasoning, and thinking: issues, challenges, and implications. In D. Ben-Zvi & J. Garfield (Eds.), *The challenges of developing statistical literacy, reasoning, and thinking*. Netherlands: Kluwer.

Gigerenzer, G., Swijtink, Z., Porter, T., Daston, L., Beatty, J., & Kruger, L. (1989). *The empire of chance: How probability changed science and everyday life*. Cambridge: Cambridge University Press.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Chicago: Aldine.

Gravemeijer, K. P. E. (1994). Educational development and developmental research in mathematics education. *Journal for Research in Mathematics Education, 25*(5), 443–471.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80*, 237–251.

Konold, C. (1989). Informal conceptions of probability. *Cognition and Instruction, 6*(1), 59-98.

Konold, C., Pollatsek, A., Well, A., Lohmeier, J., & Lipson, A. (1993a). Inconsistencies in students' reasoning about probability. *Journal for Research in Mathematics Education, 24*(5), 393-414.

Saldanha, L. A. (2003). *"Is this sample unusual?": An investigation of students exploring connections between sampling distributions and statistical inference*. Vanderbilt University, Nashville.

Saldanha, L. A., & Thompson, P. W. (2002). Conceptions of sample and their relationship to statistical inference. *Educational Studies in Mathematics, 51*(3), 257-270.

Simon, M. A. (1994). Learning mathematics and learning to teach: Learning cycles in mathematics teacher education. *Educational Studies in Mathematics, 26*(1), 71–94.

Sowder, J. T., Philipp, R. A., Armstrong, B. E., & Schapelle, B. P. (1998). *Middle-Grade Teachers ' Mathematical Knowledge and Its Relationship to Instruction: A Research Monograph*. Albany: State University of New York Press.

Steffe, L. P., & Thompson, P. W. (2000). Teaching experiment methodology: Underlying principles and essential elements. In R. Lesh & A. Kelly (Eds.), *Handbook of research design in mathematics and science education* (pp. 267-307). Hillsdale, NJ: Erlbaum.

Thompson, A. G. (1984). The relationship of teachers' conceptions of mathematics teaching to instructional practice. *Educational Studies in Mathematics, 15*, 105–127.

Thompson, A. G., & Thompson, P. W. (1996). Talking about rates conceptually, Part II: Mathematical knowledge for teaching. *Journal for Research in Mathematics Education, 27*(1), 2-24.

Thompson, P., Liu, Y., & Saldanha, L. A. (2004). *Intricacies of statistical inference and teachers' understanding of them*. Paper presented at the 33rd Carnegie Symposium on Cognition, "Thinking with data", Pittsburgh, PA.

Thompson, P. W., & Liu, Y. (2002). *What is the probability that my car is red: tensions in the development of probabilistic reasoning*. Paper presented at the Annual conference of American Educational Research Association, New Orleans.

Thompson, P. W., & Saldanha, L. A. (2000). *Epistemological analyses of mathematical ideas: A research methodology*. Paper presented at the Twenty Second Annual Meeting of the International Group for the Psychology of Mathematics Education - North American Chapter, Tucson, AZ.