PRESERVICE TEACHERS' CONCEPTIONS OF VARIATION

Katie Makar   Dan Canada

University of Queensland   Eastern Washington University

This paper reports on common findings from two recent studies of preservice teachers' conceptions of variation, one involving prospective elementary teachers and the other prospective secondary teachers. The studies both found that preservice teachers tend to use different conceptions of variation when the context of a problem changes. In addition, their descriptions of the spread of a distribution were similar, with both studies reporting the teachers using informal terminology in preference to standard descriptions (shape, mean, etc.). Implications for practice are discussed.

INTRODUCTION

Much research is being conducted about how students learn and conceptualize variation (Lee, 2003; Ben-Zvi & Garfield, 2004). Preservice teachers, however, often lack the experience and depth of understanding necessary to engage their students in the kinds of tasks that show promise in promoting the kinds of statistical understanding recommended by the research community. In particular, teachers are encouraged to provide students with authentic, inquiry-based tasks meant to develop children's reasoning about variation and distribution, but little is known about how the teachers themselves reason about variation. The purpose of this paper is to highlight two findings common to each author's doctoral research, with Canada's (2004) focus being on elementary preservice teachers and Makar's (2004) being on secondary preservice teachers. The two findings from these studies provide insight into strengths and barriers that preservice teachers have in their statistical reasoning about variation. One finding concerns their use of non-standard language when describing elements of a distribution, particularly spread. Another finding concerns the relevance of the context in which questions are embedded. The authors' studies build upon and add to previous research on conceptions of variation, most of which is aimed at students in grades 3-12. Research has provided few exploratory results on the conceptions of variation held by teachers or preservice teachers.

PREVIOUS RESEARCH

Much of the recent work on how learners develop notions of variation has come out of work at the middle school level, although researchers often find that teachers without experience with data-handling often have similar notions of randomness and variation to that of students. Teachers' conceptions of variation have previously been studied by only a handful of researchers (e.g., Mickelson & Heaton, 2003; Watson, 2001). In a study by Hammerman and Rubin (2004), inservice teachers used the software Tinkerplots to design meaningful data representations and develop reasoned arguments in a compelling context. Because of the dearth of research on this

population, we turn to research on students at various levels to gain insight into potential paths to develop teachers' conceptions of variation.

Earlier studies have looked at student understanding of variation through tasks involving data sets and graphs, sampling, and probability situations. In considering the distribution of data sets, Mellissinos (1999) noted that college students had some awareness that only looking at the center would not capture the whole picture. At the school level, researchers found that older students generally had a higher level of understanding of variation than younger students (Torok and Watson, 2000). Watson and Moritz concluded that many students from grades 3-9 did not recognize that smaller samples were "more likely to give an extreme or biased result" (2000, p. 66). Shaughnessy and Ciancetta (2001) found that by making predictions and then conducting simulations, secondary students were able to focus more attention on the variation inherent in the set of outcomes rather than just focus on the expected value for any particular outcome.

THEORETICAL FRAMEWORK AND METHODOLOGY

Two studies, one with prospective primary teachers (Canada, 2004) and the other with prospective secondary teachers (Makar, 2004) were conducted independently in one-term preservice courses at two large public universities in the United States. Both studies were designed to strengthen teachers' conceptual understanding of probability and statistics, and focused on providing the prospective teachers with multiple hands-on experiences, conducting experiments and investigations, and interpreting data in an applied context. Data collected from the studies were primarily qualitative. Pre-post tests were also given to assess conceptual understanding of statistics, particularly reasoning about variation. Subjects were interviewed to gain additional insight into their statistical reasoning as well as to investigate the ways that the teachers articulated how they were "seeing variation". Transcriptions were analysed using a grounded theory approach, allowing common themes to emerge from the teachers' descriptions and actions. The themes fall into a framework for understanding variation which was developed from prior research (Canada, 2004). The framework posits three main *aspects* of understanding variation: expecting, displaying, and interpreting variation. Within the aspect of expecting variation, we found that our results reflected the theme involving distributional reasoning. Within the aspect of displaying variation our results reflected the theme of emphasizing decisions in context, and within the aspect of interpreting variation, our results reflected the themes concerning causes of variation.

Study of Prospective Elementary Teachers

The thirty subjects in this study (24 women, 6 men) were enrolled in a ten-week preservice course at a university in the north-western United States designed to give prospective teachers a hands-on, activity-based mathematics foundation in geometry and probability and statistics. Prior to any instruction and again at the end of the course, subjects took an in-class assessment designed to elicit their understanding on

a range of topics in probability and statistics. Six subjects were selected for additional one-hour interviews outside of class before and after instruction in probability and statistics so that their conceptions of variation could be further explored. A series of activities were conducted in class specifically designed to offer opportunities to investigate and discuss variation. The activities were centered around the three realms of data and graphs, sampling, and probability situations. Take-home surveys were given after each activity.

Study of Prospective Secondary Teachers

The seventeen subjects (14 women, 3 men) in this study, all majors in mathematics or science, were enrolled in a secondary preservice course on assessment at a university in the southern United States. About half of the preservice teachers had previous coursework in statistics or had learned statistics within a mathematics or science course for their major. The study was designed to address the larger research question of how prospective teachers used the concepts of variation and distribution to support their understanding of issues of equity in testing. A subquestion related to the results reported in this paper aimed at uncovering their understanding of the concepts of distribution and variation. The course provided the prospective secondary teachers with opportunities to examine issues of equity through interpreting large-scale, school-level, and classroom-based assessment data using the statistical learning software, Fathom™. Readings in assessment and related issues of equity and high-stakes testing were assigned to deepen their contextual understanding of the data. Rather than be a course about statistics, statistical concepts were learned "as needed" as tools to investigate and gain insight into equity in assessment through the analysis of data. At the end of the course, the prospective teachers chose a topic of interest and conducted an in depth three-week data-based investigation of an issue of equity in assessment and presented their findings both in writing and as a class presentation.

ELEMENTS OF DISTRIBUTIONAL REASONING: NOTIONS OF SPREAD

In tasks involving an evaluation of a data set or a comparison of two data sets, subject responses could be categorized according to the elements of distributional reasoning used: center, range, shape, and spread. While all four elements are critical to a more sophisticated understanding of distribution, the notion of the spread of data around the center of distribution in traditional coursework is often limited to a discussion of the standard deviation or other standard statistical measures. However, just as Torok and Watson conducted a study that "successfully explored students' understanding of variation without ever employing the phrase 'standard deviation'" (2000, p. 166), so too did we find that many of our subjects used non-standard language to convey their sense of variation when reasoning about distributions of data. This section reports examples of non-standard language used by our subjects as they evaluated and compared data sets.

Prospective Elementary Teachers

A task from the post-interview showed weights for 35 different muffins bought from the same bakery, and asked what subjects thought their own (36th) muffin might weigh. The set of data for the 35 muffin weights were shown in both a boxplot and a histogram. Here are three subjects discussing the data and their own expectations:

SP: How much would I expect my muffin to weigh? Well, I'm guessing that it could be anywhere in between, somewhere around where the *bulk of this data* is. [Circling a central part of the histogram and the boxplot]

EM: Looking over here at the histogram, and that it does seem within...112 to 115.5, it seems like that seems to be a *concentration of data*. I'm going to think that it's probably going to be in the interquartile range.

DS: This one [Boxplot] you can see more...real clearly that 50 % are *really clustered* between 112 and a half and 115 and a half.

Note how SP used the phrase "bulk of this data", EM talked about the "concentration of data", and DS considered how the data was "really clustered". In other tasks, subjects referred to data presented in dot plots as being "scattered" or "bunched", which are other examples of non-standard language. All of these terms suggest a relative grouping, and they appeal to the theme of spread.

Prospective Secondary Teachers

Similar non-standard language was articulated by the prospective secondary teachers. An interview conducted at the beginning and end of the study showed the teachers two stacked dot plots of authentic student-level data from a local middle school documenting the improvement of each student from one year to the next on the state-mandated exam. The interviewer asked the prospective teachers to compare the relative improvement on the state-mandated exam for students enrolled in a test-preparation course ("Enrichment") compared to the rest of their peers to determine if the test-preparation course was effective. Results showed that only 53% mentioned the mean at the beginning of the study (82% at the end of the study), despite the fact that the means were marked on the graph shown to them; and those with previous coursework in statistics were no more likely than those without to use the means to compare the groups. Even more (67%) of the prospective teachers talked about the spread or distribution of the data at the beginning of the study (76% at the end of the study), with most articulating these concepts using non-standard descriptions similar to those found in the study with prospective primary teachers. A few excerpts from interviews at the beginning of the study are given below:

Marg: These are more *clustered* [than the other group]. So where's there's little improvement, at least it's consistent.

Brian: It seems pretty *evenly distributed* across the whole scoring range.

June: This is kind of *dispersed off* and this is like, *gathered* in the center.

Similar non-standard descriptions were used at the end of the study:

Rachel:	It's more *clumped*, down there in the non-Enrichment.
Marg:	[Enrichment] has a much *wider spread*.
Anne:	[The Enrichment] are all kind of *scattered out* almost evenly. Whereas [the non-Enrichment] are more *bunched up* together.

EXPECTATIONS OF RANDOMNESS IN DIFFERENT CONTEXTS

Consistent with the findings of Pfannkuch and Brown (1996) and Confrey and Makar (2002), both of our studies revealed that the context of the problem had a strong influence on the prospective teachers' expectation or tolerance for variation. The prospective teachers were more likely to expect random variation when they were trying to explain outcomes *a posteriori* than when they were trying to predict them. In addition, they were more tolerant of variation in problems set in probabilistic settings (e.g. dice, coin flips) and were more likely to explain variation in real-world settings with contextual explanations, particularly before instruction. Experiences with data-handling in both courses showed improved growth and stability across contexts in understanding of variation, a critical area for applying statistical reasoning in everyday contexts. Examples from each study are given below.

Preservice Elementary Teachers

In a pre-interview, subjects were given a pair of problems each involving 60 rolls of a fair die, but situated in different temporal contexts. The first question asked subjects to provide a list for the frequencies of each face of the die which might actually occur in 60 rolls. That is, how many "1s" would they expect, how many "2s", and so on. Care was taken in wording the task so that subjects knew I wanted them to imagine what might happen if they really threw the die 60 times.

RL	I'm going to say 10 for every single one.
DC	Ok. Why? Why do you think those numbers are reasonable?
RL	I think that's {10,10,10…} going to happen. Because rolling any given number is no more likely than rolling any other number. Can it happen? Sure it can happen. I think that's {10,10,10…} going to happen.

After establishing his expectation of no variation for the frequency of faces resulting from 60 tosses, RL had a different idea in the context of the second question. For the second question, subjects were asked to assume the role of a teacher who assigned 60 rolls to students as homework. Four of the students' supposed lists of results were shown, and subjects were asked to identify which lists were authentic and which were fabricated. For example, "Lee's" results were "10,10,10,10,10,10", suggesting that was what he claimed to have really obtained. In reviewing the "Lee's" list, RL said:

RL	I don't believe that [Lee] actually got that.
DC	And yet on the previous page, you said that that's what you think would really happen? [I turn the page back to look at his own list of all tens]

RL   That's…that would be the basis of my expectations. But it would be pretty funny for, in a world of imperfect scientific conditions, to see the likelihood matched so closely. It doesn't seem to account for…

At this point, RL was clearly re-examining his earlier expectations. Putting him in a position to judge results after an event had supposedly already occurred on the second question caused him to go back and change his mind on the first question:

RL   We're living in the real world, this is not going to be 10,10,10…[He's changing his own list of "10, 10, 10, 10, 10, 10" to "12, 15, 9, 11, 8, 5"]

DC   This [new list] is more like what you think might really happen?

RL   Yeah, oh yeah. The reason why I would not go for this [all tens] after all is because you're going to see a range of results. I'm changing my mind [on the first question] because I was considering average but not considering variation. You need to consider variation to get the full picture.

Based on the results of other subjects who responded similarly to RL, the context of the first question (predicting results *a priori*) has a markedly different effect on reasoning about variation than the context of the second question (evaluating results *a posteriori*). Due in part to the cognitive conflict induced by the sequencing of these two die-tossing questions in the pre-interview, subjects such as RL began a shift in expectations that led to a more consistent appreciation for variation after instruction.

Preservice Secondary Teachers

The secondary preservice teacher study examined the influence of thinking about variation in a probabilistic setting versus a real-world setting. Overall, performance on deeply contextual problems on the pretest was consistently among the greatest areas of difficulties for the secondary preservice teachers. Most of them were fairly comfortable thinking probabilistically in dice and coin problems, but tended to think deterministically if the same problem were posed in a real-world context like one might find in the media. One example comes from a pair of problems from the pre-post tests, chosen to assess the teachers' expectation of variation in small samples, with identical structure and data (adapted from Pfannkuch & Brown, 1996). One problem was set in the context of dice whereas the other problem described the number of children born with deformities in various regions in New Zealand. Pfannkuch and Brown noted that even though the problems draw on the same statistical knowledge, the context of dice encourages more tolerance of variation for small subgroups, while less variability is tolerated for real-world contexts. This result was found as well in the secondary preservice teacher study where the prospective teachers had little problem tolerating random variation in the probabilistic setting of dice, with 67% attributing variability of outcomes to randomness. However, they were likely to attribute the variability of children's deformities in the real-world example to contextual factors (83%), for example as evidence of a nearby chemical plant. The ability to tolerate random variation increased significantly after instruction

with 83% attributing variation to randomness in the probabilistic setting and 50% in the real-world setting on the post-test.

SUMMARY AND IMPLICATIONS

One goal of education is to provide students with an understanding of concepts in multiple contexts so that they can apply their understanding to complex problems encountered outside of school. Little is known about how teachers' concept of variation changes across problem contexts, even though teacher knowledge is a critical factor in student understanding. The studies reported here indicate that prospective teachers often hold competing beliefs about random variation when the setting of the problem changed. For example, many subjects considered their own preference for rain, tolerance in waiting for trains, interest in seeking justice for deformed children, or desire for weighty muffins in drawing conclusions. In these cases, the context invited the subject to consider whether or not the amount of variation shown was personally desirable or appropriate. In both studies, however, the preservice teachers' tolerance for random variation stabilized and improved across multiple contexts and settings after instruction.

Both authors were also interested in gaining insight into prospective teachers' understanding of variation. Although the pre-post tests were designed as one measure of their understanding of variation, it did not probe into *how* the preservice teachers would articulate seeing variation. The prospective teachers tended to take note of qualitative attributes of variation more often than quantitative or conventional ones (e.g., shape, center) and their choice of words were very similar in both studies.

This kind of non-standard, informal use of language use needs to be given a greater emphasis in research on statistical reasoning. Describing a distribution as "more clumped in the center" conveys a more distribution-oriented perspective than quoting standard deviation or range. Research on adults' statistical reasoning has often focused on descriptive statistics (e.g., graphical interpretation, measures of center and sometimes spread), or inferential statistics (e.g., sampling distributions, hypothesis testing). These are often the only types of statistical training offered for teachers. We would argue a need for an intermediate level of coursework, located between descriptive statistics and inferential statistics; using the results of these studies, one that develops greater sense of variation and informal inference to promote teachers broader awareness of concepts of distribution and variation. Furthermore, we believe that teachers need to develop understanding and respect for the informal language their students use when describing distributions. There are several reasons for this. For one, teachers need to learn to recognize and value informal language about concepts of variation and spread to better attend to the ways in which their students use this same language. Secondly, although the teachers in this study are using informal language, the concepts they are discussing are far from simplistic and need to be acknowledged and valued as statistical concepts. Thirdly, the scaffolding afforded by using more informal terms, ones that have meaning for the students may

then help to redirect students away from a procedural understanding of statistics and towards a stronger conceptual understanding of variation and distribution. A fourth benefit of using informal language is to broaden students' opportunities for access to statistics, an important consideration for educational equity.

References

Ben-Zvi, D. & Garfield, J. (Eds.) (2004). Special issue: Research on reasoning about variability. *Statistics education research journal 3*(2).

Canada, D. (2004). Elementary preservice teachers' understanding of variation. Unpublished doctoral dissertation. Portland State University, Portland.

Confrey, J. & Makar, K. (2002). Developing secondary teachers' statistical inquiry through immersion in high-stakes accountability data. PME-NA24, Athens, GA.

Hammerman, J. & Rubin, A. (2004). Strategies for managing statistical complexity with new software tools. *Statistics Education Research Journal 3*(2).

Lee, C. (Ed.) (2003) *Reasoning about variability: A collection of current research studies. (Proceedings of the Third International Research Forum on Statistical Reasoning, Thinking, and Literacy).* Mt. Pleasant, MI: Central Michigan University.

Makar, K. (2004). Developing statistical inquiry: Prospective secondary mathematics and science teachers' investigations of equity and fairness through analysis of accountability data. Unpublished doctoral dissertation. University of Texas at Austin.

Mellissinos-Lernhardt, M. (1999). What the mean does (and does not) tell students about a distribution. Paper presented at the Research Pre-session of the 77th Annual Meeting of the National Council of Teachers of Mathematics.

Mickelson, W. & Heaton, R. (2003). Purposeful statistical investigation merged with K-6 content: Variability, learning and teacher knowledge use in teaching. In C. Lee (Ed.) *Proceedings of the Third International Research Forum on Statistical Reasoning, Thinking, and Literacy (CD-ROM).* Mt. Pleasant, MI: Central Michigan University.

Pfannkuch, M. & Brown, C. (1996). Building on and challenging students' intuitions about probability. *Journal of Statistics Education, 4*(1).

Rubin, A., Bruce, B., & Tenney, Y. (1991). Learning about sampling: Trouble at the core of statistics. In D. Vere-Jones (Ed.), *Proceedings of the Third International Conference on Teaching Statistics, Vol. 1,* 314-319. Voorburg, The Netherlands: ISI.

Shaughnessy, J. & Ciancetta, M. (2001). Conflict between students' personal theories and actual data: The spectre of variation. Paper presented at the Second International Research Forum on Statistical Reasoning, Thinking, and Literacy. Armidale, NSW Australia.

Torok, R. & Watson, J. (2000). Development of the concept of statistical variation: An exploratory study. *Mathematical Education Research Journal, 12*(2), 147-169.

Watson, J. (2001). Profiling teachers' competence and confidence to teach particular mathematics topics: The case of chance and data. *Journal of Mathematics Teacher Education 1*(2), 1-33.

Watson, J. & Moritz, J. (2000). Developing Concepts of Sampling. *Journal for Research in Mathematics Education, 31*(1), 44-70.