MIDDLE SCHOOL STUDENTS' THINKING ABOUT VARIABILITY IN REPEATED TRIALS: A CROSS -TASK COMPARISON

J. Michael Shaughnessy, Dan Canada, and Matt Ciancetta

Portland State University

This paper summarizes the thinking of 84 middle school mathematics students' about variability in three stochastics tasks that involve repeated trial. Differences in students' acknowledgement of variability were found, depending on whether the task was from a sampling environment, or a probability environment. Students' tended to neglect variability in the probability environment. We conjecture that the way that probability is normally introduced to students is part of the cause of this phenomenon.

INTRODUCTION

Prior to several years ago there had not been much previous research focused on students understanding of variation. The concept of variability was proclaimed to be a missed opportunity in research on students' understanding of data and chance (Shaughnessy, 1997). Much of the previous work on students' understanding of data and chance has concentrated on means (e.g. Mokros & Russell, 1995) or intuitions on probabilities of outcomes and comparisons of relative likelihoods of outcomes (Fischbein & Schnarch, 1997; Konold et al, 1993). Questions about variability tend to involve possibilities for repeated outcomes from sampling, or data from repeated trials of a probability experiment, or shapes of distributions of outcomes. In this paper we report on middle school students acknowledgement of variation across three 'repeated trials' tasks.

RECENT RESEARCH

In the past four years some initial research into students' thinking about variability has begun. Shaughnessy, Watson, Moritz, & Reading (1999) found a variety of types of student thinking about variability in a repeated sampling environment. When presented with a known mixture of colored objects (say 50% red, 50% other colors), most of a sample of over 700 middle and secondary mathematics students from three countries acknowledged variability in the numbers of reds that will be obtained when repeated samples were pulled from the mixture. However, students differed in the way they presented variability in their predictions, and in their reasons for their predictions. When six samples of size ten (with replacement and mixing in between each sample) are drawn from a 50% red mixture, some students predicted a 'reasonable' spread around the expected value of 5 reds in 10 (e.g., 4,7,5,8,6,5—"because they will be around 5, but not exactly"), while others predicted 'high' (6,8,7,6,9,10—"because there are more reds") or 'wide' (4, 0, 10, 2, 9, 3—"because anything can happen"). These researchers also found that upper secondary students who had studied probability had a greater tendency to disregard variation in such predictions on sampling (5,5,5,5,5,5—"because 5 is the most likely outcome each time) than middle school or lower secondary students. Similar results have been reported in an analysis of interviews on sampling situations obtained from students aged 9 to 18 by Reading and Shaughnessy (2000). Recently, Watson has reported results of younger students thinking about variation (Watson, 2002). Watson and her colleagues have also worked on developing a scheme for describing and measuring levels of students' understanding of variability (Watson, 2000; Torok & Watson, 2000, Watson et al, to appear).

THIS RESEARCH

Each of the three tasks reported in this research involves predicting the results of repeated trials: predicting outcomes of repeated samples from a mixture; predicting the distribution of outcomes for repeated rolls of a die; and predicting the results of repeated samples of spinner trials. The questions we are interested in investigating include: 1) What differences, if any, occur in the way students predict results from repeated trials across the three tasks? 2) What reasons do students give for their predictions of outcomes on repeated-trials tasks? 3) How do their reasons differ across task environments?

PROCEDURES.

In the Fall of 2002 survey data was gathered to investigate students' acknowledgment, description of, and reasoning about variability. Tasks involving variability in three environments— sampling, probability, and data sets— were administered to over 300 students in ten classrooms from six schools, two middle schools and four secondary schools. Five of the six schools were located in a large metropolitan area of the United States (2 urban and 3 suburban schools) with the sixth school from a rural location. Each of the six schools involved in this research project has one research project class, in which we are gathering survey, individual interview, and whole class video data. Four of the schools have contributed an additional comparison class, in which only the survey data is being gathered. In this paper we will focus on some initial survey results of the middle school students' thinking about variability in the outcomes from repeated-trials tasks in the sampling and the probability environments. This research is part of a multi-year research project[1] to investigate the development of secondary and middle school students' conceptions of variability.

Eighty-four middle school students in three classes (2 Grade 6 suburban, 1 Grade 7 urban) were administered a written survey investigating their thinking about variability on tasks involving the sampling, probability, and data set environments. The three repeated trials tasks of interest for this paper are given below. We will refer to them respectively as "T1. The Sampling Task", "T2. The Dice task", and "T3. The Spinner task" for the purposes of discussion. In each task, there were several questions that preceded the ones given below to help launch the environments with the students (e.g. "How many reds would you expect to get in one sample of 10? Would it be the same everytime? What would surprise you? What is the chance the spinner lands on the shaded area on one spin? Does 1 or 6 have a better chance of being rolled, or are they the same? Why?)

T1. The Sampling Task

Suppose you have a container with 100 candies in it. 60 are red, and 40 are yellow. The candies are all mixed up in the container. You pull out a handful of 10 candies and count the number of reds.

Suppose six of your classmates did this experiment, each of them pulling out 10 candies. (After each pull, the candies are put back and remixed).

a) What do you think is likely to occur for the numbers of red candies that each classmate would pull out? (Write the numbers of reds in the spaces).

_______ _______

_______ _______

_______ _______

b) Why do you think this?

T2. The Dice Task. Consider rolling a normal six-sided die.

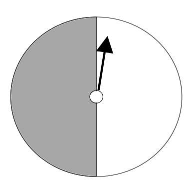
Imagine you threw a die 60 times. Fill in the table below to show how many times each number might come up. Why do you think this?

Number on Dice	How many times it might come up
1	
2	
3	
4	
5	
6	
TOTAL	60

T3. THE SPINNER TASK A CLASS USED THE SPINNER BELOW.

Suppose that you were to do 6 sets of 50 spins. Write a list that would describe what might happen for the number of spins out of 50 the spinner would land on the shaded part in each of the 6 sets of 50 spins.

_______, _______, _______, _______, _______, _______



Results and Discussion. Each student was assigned a code indicating whether they acknowledged variation for the outcomes on each task, and how they acknowledged it. Responses were coded R (Reasonable), H (High), W (Wide) L (Low) or with the numeral 6 (for T1), 10 (for T2), or 25 (for T3) if students wrote all 6's or all 10's or all 25's on a list. This type of coding is similar to ones used by previous researchers on repeated trials tasks. The number of students in each class who responded with strings of identical results, such as 6,6,6,6,6,6 for the numbers of reds in the six pulls of the Sampling Task, or 10,10,10,10,10,10 for the number of each of the outcomes from sixty rolls in the Dice Task is recorded in Table 1. For example, the entry 4 – 21 – 5 in the Grade 7 column indicates that there were 4 students who responded 6,6,6,6,6,6; 21students who responded 10,10,10,10,10,10; and 5 students who responded 25, 25, 25, 25, 25, 25 to the tasks T1, T2, and T3 respectively in that class. Also recorded in Table 1. are the numbers of students who predicted a "reasonable" spread in the outcomes for at least one of three repeated trials tasks. In the Sampling Task, outcomes with a range of ≤ 7 for the numbers of reds were considered "reasonable", while responses like 1, 7, 4, 10, 9, 0 were coded as "Wide". If all six outcomes on the list were numbers ≥ 6, the response was coded "High". If all six outcomes were numbers ≤ 6 the response was coded "Low." Similar decisions were made for the other two tasks. For example, a response list with 5 ≤ "numbers"≤ 15 for the frequencies of the die outcomes was considered a "reasonable" spread, as was a response list with a range from 15 to 35 "shaded landings" for the six sets of 50 trials of the Spinner Task.

Classes→	G7 N=29	G6 N=25	G6 N=30	Totals N=84
Tasks →	T1- T2 -T3	T1- T2 -T3	T1- T2 -T3	T1- T2 -T3
6 – 10 – 25 "No Variation" Totals[1]	4 – 21– 5	3 – 13 – 4	0 – 12 – 4	7 – 46 – 13
R – R – R "Reasonable Variation" Totals[2]	19 - 7 - 13	16 - 8 - 14	24 -10 -14	59 - 25 - 41

Table 1. Frequencies of "No Variation" and "Reasonable Variation" responses for each task in each class

1. 6-10-25 indicates the number of students who responded 6,6,6,6,6,6, or 10,10,10,10,10,10, or 25,25,25,25,25,25 respectively for the results of six trials on that task in that class.
2. R-R-R indicates the number of students who responded with a 'reasonable' spread respectively for the results of six trials on that task in that class.

Table 1. indicates that there was a very strong tendency for these students *not* to acknowledge variation when predicting the frequency distribution of outcomes for the dice problem. More than half the students predicted 10,10,10,10,10,10 for the frequencies of the six outcomes for 60 rolls of the die. On the other hand, most of the students *did* predict lists of outcomes for the repeated trials of the Sampling Task and the Spinner Task that had some sort of spread in the repeated outcomes (91% for the Sampling Task and 85% for the spinner task). Over 70% of the students predicted "Reasonable" spreads for the repeated outcomes on the Sampling Task, while only 48% predicted "Reasonable" spreads for the Spinner Task, and only 30% for the Dice Task. These results are quite

consistent across all three classes, and both grade levels. These middle school students clearly felt that the results of the Dice Task should behave quite differently than the results of the Sampling or Spinner Tasks. A comparison of the students' reasons for their decisions on the Sampling and Dice Tasks may help us to understand the differences in their thinking about the two tasks.

> Student A: (On T1) "5,6,5,4,6,7…I'd expect 6, a lil more and a lil less." (On T2) 10,10,…,10…This is reasonable since each number has 10 chances."
> Student B: (On T1) "6,7,8,5,9,4…Because there are more red". (On T2) "10, 10, …, 10…They all have an equal chance of winning."
> Student C: (On T1) "6,5,4,3,6,5…You're not always going to get 6." (On T2) "10, 10, …10, They all have an equal chance of rolling."
> Student D: (On T1) "3,4,5,6,7,8…There are more reds." (On T2) "10, 10, …,10…Each number has a one out of six chance."
> Student E: (On T1) "6,10,0,5,8,9....Students can pull any number." (On T2) " 10, 10,…,10…Each number has an equal chance."
> Student F: (On T1) "6,7,5,6,6,5…Most of the candies are red." (On T2) "10, 10, …,10…There is only one of each number so each number has the same chance."

In their responses to the Dice Task, the majority of the students were focusing only on the theoretical probability of a single outcome for one roll of the die, 1/6 for any number, whereas they were much more likely to consider a range of possible outcomes in either the Sampling Task or the Spinner Task. Previous research has indicated that the teaching of theoretical probability for single outcome events might interfere with students' attention to variability in the results of repeated trials (Shaughnessy et al, 1999). It is likely that these students have had experiences with calculating theoretical probabilities for the outcomes of rolling one or two dice in the past. They know they should expect a probability of 1/6 for any number on one toss. On the other hand, their responses on the survey also indicate that they know that the chance that the spinner lands on the shaded part on one spin is 1/2. Knowing the probability for the spinner does not cause them to predict 25 shaded spins out of 50 every time anywhere near as often as they predict 10 for each number on 60 die tosses. To these students, the die is "supposed" to come out fair. What else could one possibly mean by the word "fair die?"

Furthermore, these students were not consistent across the three task environments on their predictions for the variability in outcomes of repeated trials. Of the 84 students in the study, only 14 of them (16%) predicted reasonable (R) lists of outcomes for *all* three tasks, and only 5 of the students indicated no variation at all on *all* three of their predicted lists (predicting all 6's, all 10's *and* all 25's respectively on T1, T2, and T3). Students' responses across the three tasks were all over the place, with the most frequent coding triad being R – 10 – R for 16, about 19%, of the individual students. These 16 students expected a "reasonable" spread of results around 6 for the Sampling Task, and around 25 for the Spinner Task, but doggedly held that all 6 numbers on the die would occur 10 times.

Based on the results of this study, we conjecture that students are likely to predict constant results for repeated trials in a familiar probability situation like the Dice Task, and to neglect the issue of variability in the frequencies of individual outcomes. Although some students did acknowledge variability on the Dice Task (Student G: "12, 11, 10, 12, 9, 9…These numbers are all *around* 10"), these students were in the minority. We believe that part of the reason for such student reasoning may be the way that probability is taught in our schools. All too often we rush our students to calculating the probability of individual events or probabilities of particular outcomes, without consideration for the variation in results that can occur in actual repeated trials. We rarely give our students opportunities to develop their intuition for a likely "range of outcomes" in repeated trials situations, especially when there is a convenient probability model, like the uniform distribution for the Dice Task, to tap. We conjecture that if we want students to attend to variability *across* a variety of environments, we will need to raise *explicit* attention to variability in those environments in our work with students. This is particularly true of probability tasks, like the Dice Task. Rather than ask, "What is the probability of getting a 6" we might better ask "If we rolled the die 25 times, how many sixes do you think you would get? Now, suppose four students each rolled the die 25 times? What would the list of the numbers of sixes each of them obtained look like?" It is not just the exact probability of an outcome that is important in data and chance, but perhaps even more so, how that outcome is situated within the distribution of outcomes for an experiment, and what the "likely range" of outcomes for the experiment will be.

References

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

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

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

Reading, C. (1999). Variation in sampling. Presented at the First International Research forum on Statistical Reasoning, Thinking, & Litercy , (SRTL I), Tel-Aviv, Israel.

Reading, C. & Shaughnessy, J. M. (2000) Student perceptions of variation in a sampling situation. In T. Hakahar & M. Koyama (Eds.), Proceedings of the 24th Conference of the International Group for the Psychology of Mathematics Education (vol. 4, pp. 89 – 96) Hiroshima, Japan.

Shaughnessy, J. M. (1997). Missed opportunities in research on the teaching and learning

of data and chance. In F. Biddulph & K. Carr (Eds.), *People in Mathematics Education* (Vol. 1, pp. 6-22). Waikato, New Zealand: Mathematics Education Research Group of Australasia.

Shaughnessy, J. M., Watson, J., Moritz, J., & Reading, C. (1999, April). School mathematics students' acknowledgment of statistical variation. *NCTM Research Presession Symposium: There's More to Life than Centers*. Paper Presented at the 77th Annual NCTM Confernce, San Francisco, California.

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

Watson, J. M. (2000). *The development of school students understanding of statistical variation.* ARC project No. A000007

Watson, J. (2002). Can grade 3 students learn about variation? *Proceedings of the Sixth International Conference on Teaching Statistics*. Durban, South Africa.

Watson, J.M., Kelly, B.A., Callingham, R.A., & Shaughnessy, J.M. (in press). The measurement of school students' understanding of statistical variation. *International Journal of Mathematical Education in Science and Technology*.

1. The work reported in this paper was supported by National Science Foundation Grant # REC-0207842. All opinions, findings, conclusions, and recommendations expressed herein are those of the authors and do not necessarily reflect the views of the funder.