

PRESERVICE TEACHERS' AWARENESS OF VARIABILITY

Edward Mooney, Dhimitraq Duni, Elisha VanMeenen, Cynthia Langrall
Illinois State University, Normal, Illinois, USA
mooney@ilstu.edu

There have been changes in the recommended content for data analysis at the K–12 level in the U.S. with the publication of Common Core State Standards (2010) and Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report: A Pre-K–12 Curriculum Framework (2005). In particular, students need to understand the nature and sources of variability, which raises the question about how well teachers understand or recognize variability. The purpose of this study was to (a) examine the awareness of variability exhibited by mathematics education majors at the elementary (Grades K–5), middle (Grades 6–8) and secondary (Grades 9–12) levels and (b) to determine what that might imply about their preparation to teach data analysis. Results indicate that the preservice teachers had a limited awareness of variability even after instruction.

BACKGROUND

“Variability is at the heart of statistics” (Garfield & Ben-Zvi, 2005, p. 29) and although people may not know many of the statistical methods used to measure variability, they should recognize the prevalence of variability and know that some things vary a little and others vary a lot. Moreover, they should understand that the variability in data can often be explained (Garfield & Ben-Zvi, 2005). Recent recommendations for K–12 mathematics (e.g., GAISE Report) encourage teachers to help students understand the nature and sources of variability. For many teachers, the introduction of these ideas occurs in college during their teacher preparation program. At our university, we provide data analysis-focused courses for each level of teacher preparation – elementary, middle, and secondary. The purpose of this study was to describe students’ awareness of variability at the beginning and end of these courses and to determine what implications their level of awareness has in preparing teachers to teach ideas of variability to K–12 students.

METHODOLOGY

Participants

The participants for the study, conducted during the Spring semester of 2013, were preservice teachers who were enrolled in a mathematics content course that emphasized data analysis. The course for elementary preservice teachers addressed statistics, probability, and algebra. The course for middle level preservice teachers was a one-semester probability and statistics class designed specifically for teaching in the middle grades. Secondary level preservice teachers take two sequenced courses: a course on applied probability models and a course on statistics and data analysis. The secondary students in this study were enrolled in the latter.

Data Collection

A pretest was administered in each of the courses within the first three weeks of the 16-week semester and the same test was administered within the last two weeks of the semester. The test consisted of 18 multiple choice questions that required students to explain their selections. In total, 118 students (68 elementary, 42 middle and 8 secondary) completed both the pre- and posttests. Questions dealt with variability in data displays—dot plots, scatter plots, and boxplots; variability in sampling, and measures of variability. After the posttest, 15 students (6 elementary, 6 middle, and 3 secondary) volunteered to be interviewed. In the interview, students were reminded of their answers and asked to provide a verbal justification for how each question was answered.

Analysis

The multiple choice answers to the pre- and posttests were compiled in tables showing the percent of students who gave each response. The tables were used to examine changes in student responses. Rubrics were developed to examine the interview responses based on the four levels of consideration of variation described by Reid and Reading (2008): *strong* consideration, *developing*

consideration, *weak* consideration, or *no* consideration of variation. Once descriptors for each task were agreed upon, student responses were scored accordingly.

RESULTS

In this paper, we report the findings for two tasks. The first task, *Flipping Coins*, is similar to a spinner task used by Watson, Kelly, Callingham, and Shaughnessy (2003). Students were asked to consider the outcome if 25 students in a class each flipped a coin 50 times and recorded the number of times a “head” appeared. Students were given four graphical representations of data (See Figure 1) and asked to determine whether each display could be the result of a class experiment or whether the results were made up. Figure 2 shows the distribution of responses on the pre- and posttests. Shaded rows indicate the percentage of students that gave the ideal response for each graph. As seen in the results, students were consistent in considering Graph C as representing made up results. Growth in the percent of ideal responses occurred across the other three graphs with graph B showing the lowest percent of students giving the ideal response.

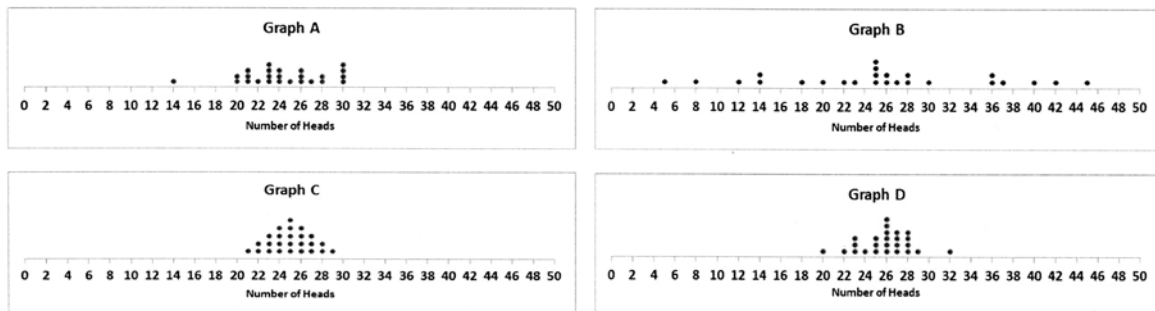


Figure 1. Graphs for *Flipping Coin* task.

Graph A	Overall (N = 118)		Elem. (N = 68)		Middle (N = 42)		Second. (N = 8)	
	PRE	POST	PRE	POST	PRE	POST	PRE	POST
Response								
Experiment	70.3%	86.4%	70.6%	86.8%	66.7%	85.7%	87.5%	87.5%
Made Up	25.4%	11.9%	25.0%	11.8%	40.5%	12.5%	12.5%	12.5%

Graph B	Overall (N = 118)		Elem. (N = 68)		Middle (N = 42)		Second. (N = 8)	
	PRE	POST	PRE	POST	PRE	POST	PRE	POST
Response								
Experiment	38.1%	28.0%	39.7%	35.3%	35.7%	16.7%	37.5%	25.0%
Made Up	54.2%	69.5%	54.4%	64.7%	54.8%	78.6%	50.0%	62.5%

Graph C	Overall (N = 118)		Elem. (N = 68)		Middle (N = 42)		Second. (N = 8)	
	PRE	POST	PRE	POST	PRE	POST	PRE	POST
Response								
Experiment	9.3%	6.8%	8.8%	5.9%	9.5%	9.5%	12.5%	0.0%
Made Up	84.7%	86.4%	83.8%	86.8%	85.7%	85.7%	87.5%	87.5%

Graph D	Overall (N = 118)		Elem. (N = 68)		Middle (N = 42)		Second. (N = 8)	
	PRE	POST	PRE	POST	PRE	POST	PRE	POST
Response								
Experiment	76.3%	83.1%	77.9%	83.8%	71.4%	78.6%	87.5%	100.0%
Made Up	18.6%	6.8%	17.6%	5.9%	21.4%	9.5%	12.5%	0.0%

Figure 2. Results of the pretest and posttest for the *Flipping Coins* task. Percentages for *No Response* and *Not Sure* were omitted.

Interview responses revealed that students were able to display understanding based on four levels of consideration of variability. At the highest level, two of 15 students demonstrated a *strong* consideration of variability by giving a specific expectation of variability in relation to a specific measure of center. For example, Charlie stated that “The expected value of the number of heads is 25, so a lot of them are at about 25...Most of them should be within this area of 20 to 30.”

At the next level, six students showed a *developing* consideration of variability by giving an explanation that indicated a vague expectation of variability. Some consideration of variability was evident, but specific measures of variability or a reasonable range for the data were not

mentioned. Carla, in her response, stated that “I knew that the average perfectly would be like 25, so all the numbers were close to there.”

A *weak* consideration of variability can be characterized by the absence of exact measures and a focus on either center or spread, but with no connection between the two. Examples of the six student responses at this level included the following: “[The data] varied and they did not seem to follow a pattern.” “This one is not spread enough.” “Sometimes it will be a little bit more and sometimes a little less, but it is always going to be pretty close to somewhere in the middle.”

No consideration of variability was recorded for one student response. The response showed that the student focused mainly on parts of the graph that did not relate to center or spread. “I just felt like it did look like a class experiment even though there was some in the front of the pack and it looked like it was spread out in the middle.”

The second task, *Container of Candy* (see Figure 3), is based on a task developed by Torok and Watson (2000). Figure 4 shows the distribution of responses from the pre- and posttests. The shaded rows show the percentage of students providing the ideal response. As can be seen in the results, little change occurred in the percent of students selecting the ideal response.

Container of Candy
 Suppose you have a container with 100 candies in it. 50 are red, 20 are yellow, and 30 are green. The candies are all mixed up in the container. You pull out 10 candies and record the number of red candies drawn. You then put the candies back and mix up the candies. You draw 10 candies a total of 6 times replacing the candies before mixing them and drawing again.

Which one of these lists do you think best describes the result you would get? **Explain your thinking.**

- a. 5, 9, 7, 6, 8, 7
- b. 3, 7, 5, 8, 5, 4
- c. 5, 5, 5, 5, 5, 5
- d. 2, 3, 4, 3, 4, 4
- e. 10, 10, 10, 10, 10, 10
- f. Not Sure

Figure 3. The *Container of Candy* task.

Response	Overall (N = 118)		Elem. (N = 68)		Middle (N = 42)		Second. (N = 8)	
	PRE	POST	PRE	POST	PRE	POST	PRE	POST
5, 9, 7, 6, 8, 7	18.6%	22.9%	20.6%	25.0%	16.7%	16.7%	12.5%	37.5%
3, 7, 5, 8, 5, 4	56.8%	58.5%	45.6%	48.5%	73.8%	73.8%	62.5%	62.5%
5, 5, 5, 5, 5, 5	14.4%	14.4%	23.5%	19.1%	0.0%	9.5%	12.5%	0.0%
2, 3, 4, 3, 4, 4	4.2%	1.7%	2.9%	2.9%	7.1%	0.0%	0.0%	0.0%
All 10's	4.2%	0.8%	4.4%	1.5%	2.4%	0.0%	12.5%	0.0%

Figure 4. Results of the pretest and posttest for the *Container of Candy* task. Percentages for *No Response* and *Not Sure* were omitted.

Two of the 15 interview responses were classified as having a *strong* consideration of variability. At this level, students provided a clear range for the number of candies to be drawn. Karla said, “I chose [B] because 3 is 2 less than 5, and 7 is 2 more than 5 so the numbers are at 5 now, almost like the standard deviation.” Loren, in her justification for selecting B, stated “You might get a couple of 5’s or you might get some lower like 3 or 4 and some higher like 7 or 8.”

Responses at the *developing* consideration level typically included comments that some results should be around 5. Jennie stated, “I thought [B] was the most diverse but still around [5]. You are not going to get 10 every time; you are not going to get 5 every time. You sometimes will get 5; sometimes you will not. Sometimes you will get somewhere close to 5.” Six student responses were classified at this level.

Two student responses were considered to show a *weak* expectation of variability. These students were aware that not all outcomes would be 5 but they were unclear of the expectation. In explaining why she chose B, Jane stated “The probability of picking all reds is going to be in the 5

range because half of them are reds and A had a lot more larger number, the 8 and 9, and I felt that was a lot of reds and then C and E had all exactly the same number of reds [that is] probably not going to happen. And, D had a lot of lower number of reds which is not as likely to happen...so I chose B because there was a good representation of all of it.” At the lowest level, *no* consideration of variation, responses showed no expectation in changes in the number of red candies drawn. Typical of the five student responses, Mary stated, “I thought that since you have the greatest chance of getting red and you are putting them back each time, you have about half the time you get red and half the time you get another color, so I just picked the 5’s [option C].”

CONCLUSION

The results of the two tasks highlight similar issues regarding students’ expectation of variability. Overall, students recognized that a certain amount of variability should be expected but they were less certain about how much variability would be likely to occur. In fact, the responses given by the preservice teachers were similar to responses given by middle level students in Watson’s studies (Watson, 2006). Watson (2006) reported that in the *Spinners* task, parallel to the *Flipping Coins* task, students studied were most successful in recognizing a graph similar to Graph C (See Figure 1) as being made up, and least successful at recognizing a graph similar to Graph B as being made up. The students justifications were similar to those given by the preservice teachers in this study. Watson also reported that approximately half of the middle level students in her study could successfully select the appropriate outcome for the number of candies drawn when answering a task similar to the *Container of Candies* task, which was comparable to our findings with college students.

If we compare the posttest responses of the preservice teachers across the three levels, we see that the results are similar for both tasks. The level of mathematics or type of preservice program did not seem to affect students’ answer choices. Also, levels of consideration of variability from the interview responses did not necessarily correspond to the preservice teachers’ program. That is, secondary preservice teachers did not always give responses at the highest levels of consideration and elementary preservice teachers did not always give low level responses.

The implication for instruction seems to be that students need more experiences collecting and examining multiple samples to draw patterns or conclusions about the expected results. For example, the tasks used in this study could be activities done in class with or without the support of technology. Furthermore, students need the opportunity to think about both the center and the spread, how they relate to each other, and how they relate to the outcomes seen or expected. Students frequently discussed the data with regard to the center but had difficulty expressing the idea of variability. It was rare that a student incorporated both ideas during the interview. It seems that we are building our students’ expectation that variability exists in data, but we need to provide them with experiences to help them see that a certain type or degree of variability can be expected.

REFERENCES

- Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2007). *Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report: A Pre-K–12 Curriculum Framework*. Alexandria: American Statistical Association.
- Garfield, J., & Ben-Zvi, D. (2005). A Framework for Teaching and Assessing Reasoning about Variability. *Statistics Education Research Journal*, 4(1), 92–99.
- National Governors Association for Best Practices, & Council of Chief State School Officers (2010). *Common Core State Standards for Mathematics*. www.corestandards.org/.
- Reid, J., & Reading, C. (2008). Measuring the Development of Students’ Consideration of Variation. *Statistics Education Research Journal*, 7(1), 40–59.
- Torok, R., & Watson, J. (2000). Development of the Concept of Statistical Variation: An Exploratory Study. *Mathematics Education Research Journal*, 12, 147–169.
- Watson, J. (2006). *Statistical Literacy at School: Growth and Goals*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Watson, J., Kelly, B., Callingham, R., & Shaughnessy, J. (2003). The Measurement of School Students’ Understanding of Variation. *International Journal of Mathematical Education in Science and Technology*, 34, 1–29.