



# The Gumball Machine: Linking Research and Practice About the Concept of Variability

M. Alejandra Sorto and Alexander White  
Texas State University

*Journal of Statistics Education* Volume 16, Number 2 (2008), [www.amstat.org/publications/jse/v16n2/white.html](http://www.amstat.org/publications/jse/v16n2/white.html)

Copyright © 2008 by M. Alejandra Sorto and Alexander White all rights reserved. This text may be freely shared among individuals, but it may not be republished in any medium without express written consent from the authors and advance notification of the editor.

---

**Key Words:** Variability; Sampling; Lollie Task.

## Abstract

Inspired by the research of Reading and Shaughnessy (2004), we modified an existing lesson from the National Council of Teachers of Mathematics. This lesson, a variation on the "Gumball Task", gives students the opportunity to explore and discuss the variation which occurs in sampling. This paper describes our experience using this lesson as an enrichment activity in a fifth grade classroom.

## 1. Introduction

"Imagine that a gumball machine contains 1 red, 2 green, 3 yellow, and 4 blue gumballs that were thoroughly mixed before they were put into the machine. If you got one gumball from the machine, which color do you think would come out?" is the first discussion question in a lesson from the *Navigating through Probability in Grade 6 – 8 Series* (NCTM 2003, pp. 119-120) called "Gumball Machine". Students are then asked to predict how many of each color would come out if we sampled 10 gumballs with replacement. Recent research about student understanding of variability (Reading and Shaughnessy 2004) indicates that Grade 4 students consistently give "point value" answers for similar questions not giving consideration to a "range of values" which would occur on repeated sampling. Work by researchers in Australia, New Zealand and the United States suggests that asking students to describe what would happen with repeated samples of 10 gumballs yields a more complete picture of student understanding of sampling and sampling variation.

So when asked to give a special enrichment lesson to a local fifth grade classroom, we decided to use what we learned from this research to modify the "Gumball Machine" activity (see [Appendix A](#)) with the goal of giving the students an opportunity to experience and discuss the variation that occurs in repeated sampling. This article gives a brief review of the research on the problem, describes how the activity was modified, carried out and how the students reasoned about variability in this particular setting.

## 2. What Research Says about Gumball Tasks

A summary of the research on this kind of gumball tasks (sometimes referred to as candy tasks in the U.S. or lollie tasks in Australia) is presented in a chapter by [Reading and Shaughnessy \(2004\)](#) in the book *The Challenge of Developing Statistical Literacy, Reasoning and Thinking*. [Zawojewski and Shaughnessy \(2000\)](#) analyzed a gumball task which appeared in the Grade 4 1996 National Assessment of Educational Progress (NAEP). Students were told the number of red, yellow and blue colored gumballs in a bowl and then asked how many red gumballs they would expect to get in a handful of 10. Only 1 of the 232 student responses considered actually gave a range of possible numbers for the number of red gumballs. Reading and Shaughnessy suggest that although this result is not surprising given the most frequent types of questions posed in the classrooms and textbooks, it is troubling that students may not be recognizing the role variability plays in sampling. In other studies ([Shaughnessy, Watson, Moritz, and Reading 1999](#), [Shaughnessy, Ciancetta, and Canada 2004](#)), gumball tasks were given to 400 students in Australia, New Zealand and the United States and to over 700 secondary students in the United States. In this task, however, students were also asked to predict the number of reds for six handfuls of candy. The 1999 study found a tendency for students to give too wide a prediction of the range of the possible numbers of red gumballs. But more importantly, a proportion of the students changed their minds and produced a more normative response after doing the experiment and seeing the variation. The 2004 study found similar results and additionally categorized students' responses into three broad types of reasoning: additive, proportional and distributional. Additive responses focused on the number of red gumballs in the population rather than the proportion. Proportional responses implicitly suggested or explicitly mentioned that they used the population proportion to make their predictions. Distributional responses integrated both centers and variation about their centers. Looking more broadly at reasoning about variation, [Reading and Shaughnessy \(2004\)](#) describe two main aspects: how spread out the numbers are (extreme values and the range) and what is happening with numbers contained within that range. They suggest that bringing these aspects together, as can occur with the gumball task, leads to a discussion of deviation and eventually to the concept of distribution.

### 3. Our Gumball Activity

The original activity (tasks 1-2, 4-7, 9-10 of [Appendix A](#)) had two goals: to understand that events associated with small probability can and do occur, and to realize that the distribution of data from small samples often does not reflect the parent distribution. Our modified activity added two significant goals: to take variation into account when describing possible sampling outcomes and to recognize an effect of sample size on sampling variation. This was accomplished by adding three key elements:

1. Since the original questions in the activity lead students to focus more on point estimates than variation, we added question 3 where students were asked to imagine that 5 students select 10 gumballs each and to predict the number of blue gumballs the students get (see [Figure 1](#)).
2. To see the effect of sample size, students were asked to find and plot the experimental probability for both the sample of  $n=10$  gumballs and  $n=20$  gumballs.
3. To help visualize variability across many samples, students were given circular colored stickers and asked to plot their answers for each color and sample size on large number lines in the front of the class.

These simple modifications gave students the opportunity to explore the variation that occurs from repeated sampling and note any effect of changing the sample size from 10 to 20.

#### 3.1 Launching the Activity

Students were given a copy of the modified "Gumball Machine" activity. As one of the students read out loud the first question, we were demonstrating the gumball machine and the gumballs using a cup and color chips. Then, students were asked to individually answer the question and explain their reasoning in writing. After a few minutes, students shared their responses. They all said that a blue gumball would come out. One student said, "because it has the most in the machine"; another said, "because that color takes up most of the gumballs". Even though "blue" is the best prediction for a single outcome we wanted to see if students were aware of the complement of this event. As a follow-up question, students were then asked, "what are the chances of getting a color different than blue?" and "are the chances of getting another color different than blue the same as getting blue?". As a group, they all agreed that there are more chances of getting a color different than blue "because there are 6 gumballs that are not blue compared to 4 blue gumballs out of 10".

Students moved on to the second question, "Suppose that after each gumball comes out, the gumball wizard magically puts another one of the same color into the machine so that the gumball machine always has the same number of each color of gumball. If you took 10 gumballs out of this magic machine, with the gumball wizard replacing your gumball each time, how many times do you think each color would come out?" Since the wording of the question may lead students to think that all colors occur with equal frequency instead of in the 4:3:2:1 proportion, we acted out the selection and replacement process in the front of the class. The answers to this question vary much more than the answers to the first question. Students showed a large variety of reasoning. For example, a few students said that "you don't know" or "your never know" because "it will be restored each time" or "because you take one out and your putting them back in the machine". This seems to indicate an "equiprobability bias" ([Lecoutre 1992](#)). In other words the belief that since we are selecting randomly: anything can happen and everything has the same chance of occurring. Another pair of students said "at least all once or twice" or "at least twice" because "there is ten and it's almost like pulling each one time" or "each gumball can come out equally". About a quarter of the students gave different distribution of colors such as "3 blue, 4 yellow, 1 green, 2 red," or "6 blue, 2 yellow, 1 green, 1 red" including some whose predicted outcomes contained less than 10 total gumballs. Just under two thirds of the students predicted "4 blue, 3 yellow, 2 green, 1 red" and their explanations were based mainly on the original distribution of the colors in the gumball machine.

A third launching task, our first modification to the original activity, focused only on the number of blue gumballs (see [Figure 1](#)). Phrasing the prediction question in this fashion, gave the students the opportunity to explore and discuss the variability we would expect in sampling.

Name: \_\_\_\_\_

Imagine that 5 students select 10 gumballs with replacement from the magic gumball machine with 4 blue, 3 yellow, 2 green and 1 red gumballs. Predict the numbers of blue gumballs the students get.

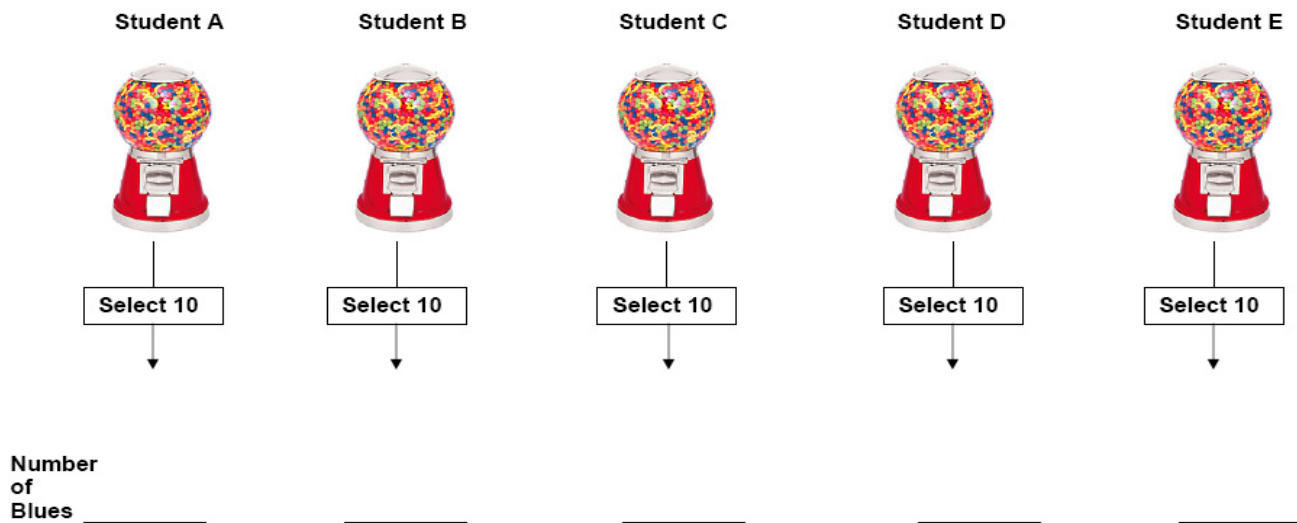


Figure 1. Third task given to students

Students' predictions for the numbers of blue gumballs varied. Most students provided a sequence of number of blue gumballs varying from 1 to 9 (though no one included both 1 and 9). Using statistical theory as a guide, we characterized the predictions as "low", "high", "wide" and "narrow". Four out of 18 students gave low predictions (with means less than 3) and one student gave a high prediction with a mean of 6. Three students wrote a "narrow" sequence of repeated numbers (e.g. "6, 6, 6, 6, 6"), suggesting no variability in the sampling results. Only one student's sequence was "wide" with a range of 6. These results indicate that the students considered here may have entered with a slightly better feel for variability than those reported in [Shaughnessy et al. \(1999\)](#). However, in the classroom discussion which followed Task 3, few students were able to articulate a justification for their choice of numbers. Those that did provide a justification did not express confidence in the correctness of their response.

### 3.2 Exploring the Outcomes

After the discussion about the prediction of the outcomes, students actually did the experiment. Students were organized in teams of two, each team was handed a cup to simulate the gumball machine and 10 color chips to simulate the gumballs. Students were asked to follow the directions on the activity sheet. After students recorded their outcomes, they were confronted with their predictions from Task 2. Most students were surprised to see that their predictions were, in their own words, "way off".

Students were asked to repeat the experiment 10 more times and then to combine the two set of data. So now students had outcomes for 10 and 20 trials. In the following tasks, the concept of "experimental probability"<sup>1</sup> was introduced by asking the students to compute the fraction of blue, yellow, green, and red they got for both 10 and 20 trials. To visualize these quantities, students were asked to place the observed experimental probabilities for each color on separate number lines (tasks 8 and 10 on the activity sheet).

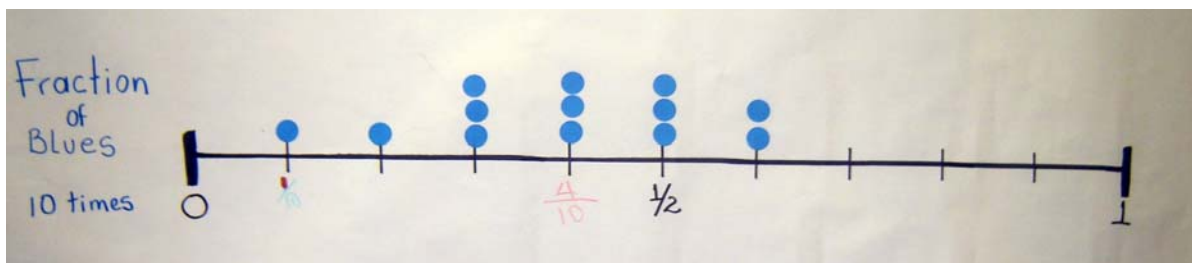


Figure 2. Dot plot of observed experimental probabilities

To show the variability of the sample proportion of blues, we surveyed the class about the observed experimental probability and asked them to place a blue dot sticker in a pre-made large number line (see [Figure 2](#)). Students were asked "Did anyone get 4 blue?" and "Where do you see this on the graph?" Students then reflected on their predictions about the number of blues for the 5 imaginary students (Figure 1). This modification of the original activity permitted a discussion why it is not likely to get 4 every time, but it is also not likely to get 0 or 10. As

recommended in [Reading and Shaughnessy \(2004\)](#), the discussion focused not only on the center of the resulting distribution but the spread and extreme values as well; setting the basis for future, more formal, reasoning about variation. This was repeated for the combined sample of 20 draws for the entire class. Comparing the two results we were able to discuss that with a larger sample you get a better representation of the parent population.

To demonstrate to the class that with an even bigger sample we can get closer to the theoretical probabilities, we asked students to report their results based on the 20 trials in the table posted in front of the classroom. Students were asked to compute the experimental probabilities for the entire class. In our case, as expected, we got numbers very close to .4-.3-.2-.1.

#### 4. Beyond the Activity

While the time allotted for enrichment activity did not permit any further investigation, there are important features we would recommend as follow up. As is, the gumball activity is an investigation in probability. To make the transition to statistical reasoning, we recommend taking out a "gumball machine" with an unknown composition of colors, having the students take a sample of 10 gumballs, and asking the students to estimate the fraction of blue balls in the machine based on the fraction observed in the sample. Following up with a discussion of how confident they are in the answer and asking what could be done to get an estimate in which we were more confident. This would lay the foundation for future discussions of informal statistical inference and the role of sample size.

Finally, as a post activity assessment, we recommend giving the students the same set up for the gumball task and then asking the following questions

1. If each student selected 10 gumballs. Which is more likely to be true?
  - a. All of the students will have exactly 1 red ball.
  - b. Most of the students will have close to 1 red ball.
2. If each student selected 10 gumballs. Which is more likely to be true?
  - a. No students will have 2 red balls.
  - b. Some, but not many students will have 2 red gumballs.

#### 5. Conclusion

Translating results from statistical education research to the classroom may seem like a daunting task. Visions of learning new software, completely revamping the curriculum or drastically changing the style of teaching may scare away teachers eager to help their students learn statistics. However, as this activity shows, by making modest modifications of an existing activity often we can implement important lessons learned from the research without too much effort. Furthermore, we can improve student understanding of variation, learn more about how our students reason about probability and statistics and perhaps in a small way add to the body of knowledge in statistics education.

#### Footnotes

<sup>1</sup> In school mathematics materials, including the original NCTM activity, empirical probability is commonly referred to as experimental probability.

#### Appendix A

### Gumball Machine

Names: \_\_\_\_\_

\_\_\_\_\_

1. Imagine that a gumball machine contains 1 red, 2 green, 3 yellow and 4 blue gumballs that were thoroughly mixed before they were put into the machine. If you got one gumball from the machine, which color do you think would come out? \_\_\_\_\_ Explain your answer.
2. Suppose that after each gumball comes out, the gumball wizard magically puts another one of the same color into the machine so that the gumball machine always has the same number of each color of gumball. If you took 10 gumballs out of this magic machine, with the gumball wizard replacing your gumball each time, how many times do you think each color would come out? \_\_\_\_\_ Explain your answer.

3. Imagine that 5 students select 10 gumballs each from the magic gumball machine with 4 blue, 3 yellow, 2 green and 1 red gumballs. Predict the numbers of blue gumballs the students get. Write your predictions on the attached answer sheet.

4. Do the following experiment to simulate the gumball machine:

- Put 4 blue chips, 3 yellow chips, 2 green chips, and 1 red chip in a paper cup.
- Mix up the chips thoroughly.
- Draw one chip out (without looking), tally its color in the middle row of the table below, replace the chip in the cup, and mix the chips again.
- Repeat these steps until you have tallied 10 trials.

Color of Chip	Blue	Yellow	Green	Red
Tally				
Number				

5. How close were your results to the predictions you made in Task 2? \_\_\_\_\_ Explain.

6. Do the experiment 10 more times, and record your data below.

Color of Chip	Blue	Yellow	Green	Red
Tally				
Number				

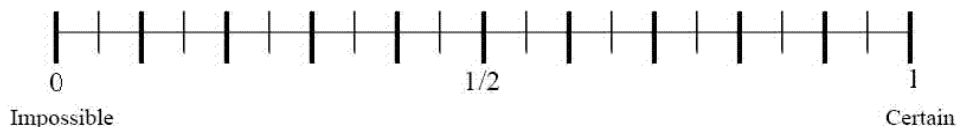
7. Combine the two sets in the table below.

Color of Chip	Blue	Yellow	Green	Red
Tally				
Number				

8. Using the table in Task 4, compute the experimental probability of drawing each color. In other words, compute the fraction of blue, yellow, green and red you got:

Color of Chip	Blue	Yellow	Green	Red
Experimental				
Probability				

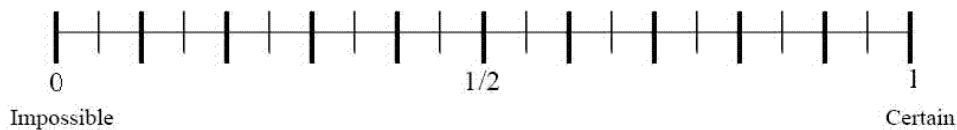
9. Use B, Y, G and R to represent the four colors. On the scale below, show the experimental probability of drawing each color you got in Task 8.



10. Now use the table in Task 7 to compute the experimental probability of drawing each color. In other words, compute the fraction of blue, yellow, green and red you got:

Color of Chip	Blue	Yellow	Green	Red
Experimental				
Probability				

11. Use B, Y, G and R to represent the four colors. On the scale below, show the experimental probability of drawing each color you got in Task 7.



## References

Lecoutre, M. P. (1992), "Cognitive models and problem spaces in 'purely random' situations," *Educational Studies in Mathematics*, 23, 557-568.

National Council of Teachers of Mathematics (2003), *Navigating through Probability in Grade 6 – 8 Series*, Reston, VA: Author.

Reading, C., and Shaughnessy, J.M. (2004), "Reasoning about variation," in *The Challenge of Developing Statistical Literacy, Reasoning and Thinking*, eds. D. Ben-Zvi and J. Garfield, Dordrecht, The Netherlands: Kluwer Academic Publishers, pp. 201-226.

Shaughnessy, J.M., Ciancetta, M., and Canada, D. (2004), "Types of Students Reasoning on Sampling Tasks," *Proceedings of the 28<sup>th</sup> Conference of the International Group for the Psychology of Mathematics Education*, 2004 Vol. 4, pp. 177 – 184. Retrieved April 17, 2008 from [www.emis.de/proceedings/PME28/RR/RR045\\_Shaughnessy.pdf](http://www.emis.de/proceedings/PME28/RR/RR045_Shaughnessy.pdf).

Shaughnessy, J. M., Watson, J., Mortiz, J., and Reading, C. (1999), "School mathematics students' acknowledgement of statistical variation," NCTM Research Pre-session Symposium: *There's More to Life than Centers*. Paper presented at the 77<sup>th</sup> Annual NCTM Conference, San Francisco, California.

Zawojewski, J. S., and Shaughnessy, J. M. (2000), "Data and chance," in *Results from the Seventh Mathematics Assessment of the National Assessment of Educational Progress*, eds. E. Silver and P. A. Kenney, Reston, VA: NCTM, pp. 235-268.

M. Alejandra Sorto  
 Dept of Mathematics  
 Texas State University  
 San Marcos, Texas 78666  
 U.S.A.  
[sorto@txstate.edu](mailto:sorto@txstate.edu)

Alexander White  
 Dept of Mathematics  
 Texas State University  
 San Marcos, Texas 78666  
 U.S.A.  
[aw22@txstate.edu](mailto:aw22@txstate.edu)