ASA AMERICAN STATISTICAL ASSOCIATION
Promoting the Practice and Profession of Statistics

NCTM NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS

The Statistics Teacher Network

Number 83 | The Statistics Teacher Network | ASA/NCTM Joint Committee on Curriculum in Statistics and Probability | Spring 2014

The Editor's Corner

Welcome to Issue 83 of the *Statistics Teacher Network (STN)*. The three articles in this issue will give you some valuable ideas for helping students create statistical displays and understand statistical concepts. The four associate editors and I have worked with the authors to provide you this timely and useful information. Our first article, from members of the special task force of the Statistical Graphics Section of the American Statistical Association, describes what they observed in the 2013 Statistics Poster Competition. You'll find examples of winning posters and learn what not to do when creating pie charts, bar charts, and graphs for comparisons. Also, there is a list of additional resources.

In the article, "My Experience Explaining Biostatistics to Elementary School Children," David Kerr from Axio Research shares what happened when he visited his daughter's elementary classroom and offers suggestions for how a teacher might duplicate the mock hands-on clinical trial.

Tim Jacobbe, associate professor, and Douglas Whitaker, a doctoral fellow, both from the University of Florida, write about a current NSF-funded Levels of Conceptual Understanding in Statistics (LOCUS) project, which focuses on developing statistical assessments in the spirit of the GAISE framework. These assessments emphasize conceptual (rather than procedural) understanding and are scheduled to be available in August.

In addition, we've included a STEW lesson, by Alexander White, M. Alejandra Sorto, Rini Oktavia from Texas State University - San Marcos. The lesson focuses on exploring variation that occurs in sampling. Students are asked questions about the probabilities related to a gumball machine which contains 1 red, 2 green, 3 yellow, and 4 blue gumballs that were thoroughly mixed before they were put into the machine. This would be a great lesson to help your students understand an important concept in statistics!

As always, I think you will find all three articles worth the read! Please continue to contact me with your ideas for improving *STN*, suggestions for articles, new teaching techniques, and/or upcoming events relevant to our cause. Please email me directly at *rpierce@bsu.edu*.

Best Regards,
Rebecca Pierce, Editor, Ball State University

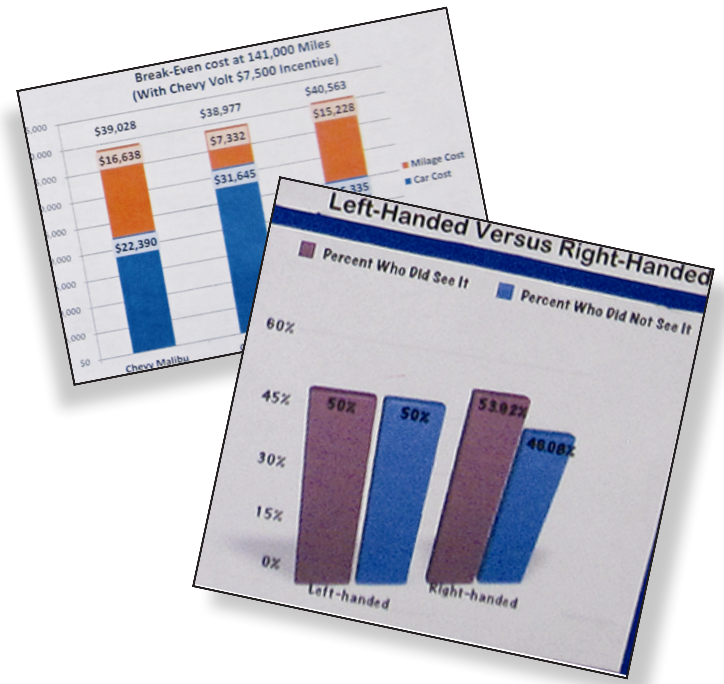
Associate Editors

Jessica Cohen – Western Washington University

David Thiel – Clark County School District, Retired

Angela Walmsley – Northeastern University, Seattle Graduate Campus

Derek Webb – Bemidji State University



Observations from the Winners of the 2013 Statistics Poster Competition—Praise and Future Improvements
p. 2

In this issue ...

Articles

Observations from the Winners of the 2013 Statistics Poster Competition—Praise and Future Improvements 2

My Experience Explaining Biostatistics to Elementary School Children 6

Lessons from the LOCUS Assessments (Part 1): Comparing Groups 13

Featured STEW Lesson Plan: The Gumball Machine 16

Announcements 23

Observations from the Winners of the 2013 Statistics Poster Competition—Praise and Future Improvements

Jürgen Symanzik, Naomi B. Robbins, and Richard M. Heiberger

The Statistics Poster Competition (*www.amstat.org/education/posterprojects*) is an annual competition jointly organized by the American Statistical Association (ASA) and National Council of Teachers of Mathematics (NCTM). We were impressed by the variety and quality of statistical graphs produced by these young researchers. In some rare cases, however, we thought alternative representation would have helped to further improve the overall message being communicated.

From Good Graphs to Even Better Graphs

We generally commend the creators of these graphs for adding titles, labels, and legends. Very well done!

Pie Charts

Pie charts can be found almost everywhere—on the web, in newspapers, and in several of the winning posters. However, there exist a few rules that should be followed when creating them.

In general, three-dimensional (3D) graphs are misleading, as areas in the front appear to be much larger than areas in the back. This is particularly true when a rim is plotted as in Figure 1 (top). At first glance, the green area in the front right appears to be much bigger than the blue area in the rear right. Therefore, two-dimensional (2D) pie charts are preferred. In addition, the order in which the individual categories are arranged in the slices of the pie chart is important.

An ordered pie chart such as in Figure 1 (bottom) makes it much easier to compare categories than an unordered pie chart such as in Figure 1 (center). There exist a few rules for best ordering a pie chart. One of those rules has been followed in Figure 1 (bottom): Start on top at the noon position and fill in the categories in a clockwise direction, starting with the largest percentage and ending with the smallest percentage. In case of an "other" category, this is often filled in last. This makes it relatively easy to see which categories jointly make up 25%, 50%, and 75% of the data. Moreover, due to the sorting, it is immediately clear which category contains a (slightly) larger percentage of the data. This is not always obvious in an unsorted pie chart, such as the one in Figure 1 (center), where it is not clear whether the orange or the blue area is bigger.

(Stacked) Bar Charts

Bar charts, including stacked bar charts, are another form of a widely used graph. However, there exist a few recommendations for making bar charts more effective.

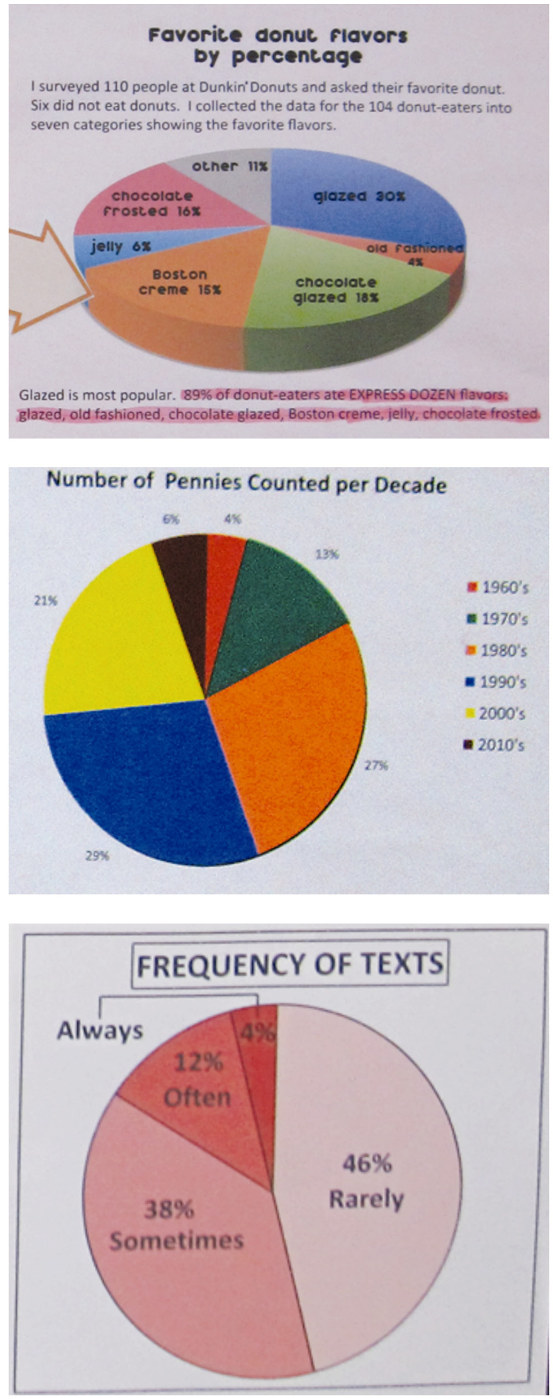


Figure 1: Three pie charts: (top) unsorted 3D pie chart, (center) unsorted 2D pie chart, and (bottom) sorted 2D pie chart

Stacked bar charts can be useful when comparing percentages. However, they are difficult to interpret if we look at totals, such as counts or a monetary value like in Figure 2 (top). In this graph, it is obvious that the right most stacked bar is higher than the left most stacked bar. But is this only due to the larger amount of the blue area at the bottom? What can we say about the orange area? Is this also larger in the right most bar (compared to the left most bar), about the same, or even smaller?

Instead of using stacked bar charts, side-by-side bar charts are preferred in many situations. Similar to 3D pie charts, we also should avoid 3D bar charts, such as the one shown in Figure 2 (center). Here, the lack of a common baseline makes it difficult to visually compare the lengths of the four bars. Typically, we only use the lengths of the bars (and not their area) when we compare multiple bars in a bar chart. This is difficult when the bars do not start on a common horizontal baseline. About how much smaller is the right most blue bar in this graph, compared to the left most purple bar? The 3D effect makes it impossible to answer this question visually (without looking at the printed numbers).

Often, 2D side-by-side bar charts are most effective, such as the one in Figure 2 (bottom). Such side-by-side bar charts are even more effective when the bars are sorted from highest to lowest (or vice versa) by the most important variable. Here, this seems to be the language students want to learn (shown in green). Due to the sorting, it is immediately evident that about three times as many students want to learn Spanish as those who want to learn the second-ranked language, which is French. Side-by-side bar charts also allow us to spot unusual patterns relatively quickly: More students know sign language (five) than students who want to learn it (three).

It does not hurt to remind readers that we technically have to compare the size of the areas in bar charts. But, as mentioned above, this is simplified to comparing the lengths of the bars (given they are all equally wide). Therefore, it is a must that all bars start at zero, since all lengths start at zero. Starting bars at values other than zero is misleading to the viewer. This rule was generally followed in the posters.

We do not want to go into an in-depth discussion of the use of colors; however, the reader should be reminded that colors may appear differently under different lighting and shading conditions, when projected onto a wall or seen on a computer screen, or when drawn on different types of paper. Therefore, some readers may have problems seeing the yellow bars in Figure 2 (bottom) while these bars are easily discernible for other readers. Colors should always be chosen with this in mind.

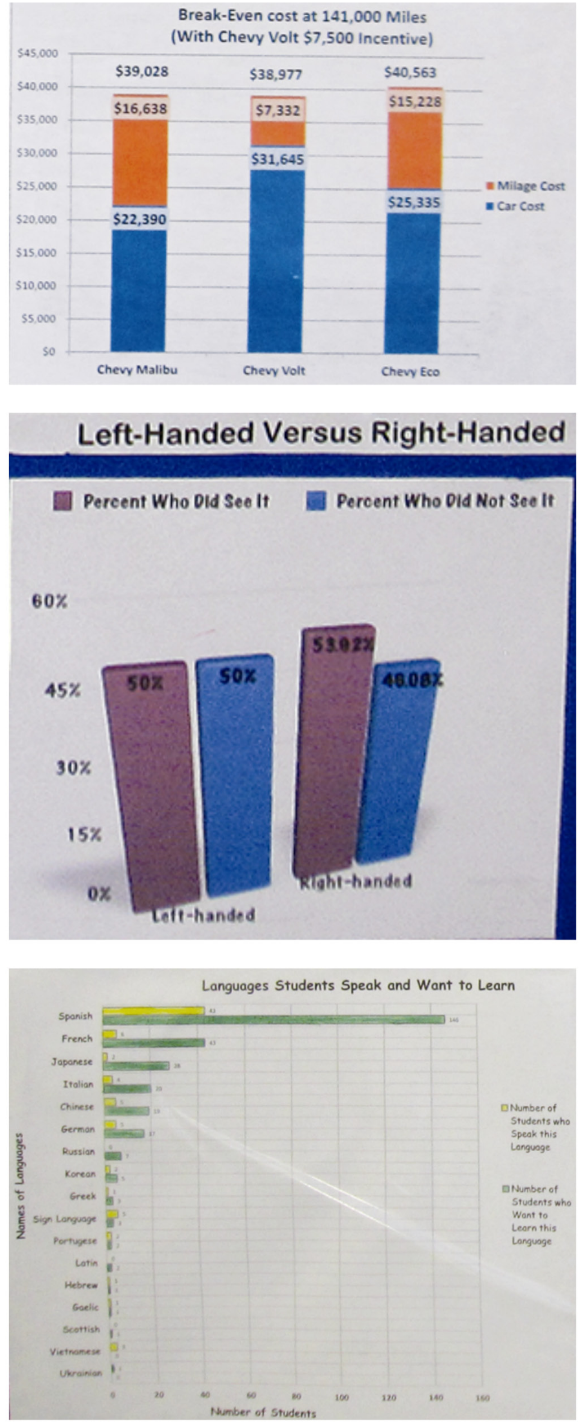


Figure 2: Three bar charts: (top) stacked bar chart, (center) 3D side-by-side bar chart, and (bottom) sorted side-by-side bar chart

Graphs for Comparisons

It is often necessary to compare data from different groups or different years. A few suitable graphs are mentioned here. In Figure 3 (top), side-by-side boxplots of a quantitative variable (the time until clothes are dry) are shown for two levels of a categorical variable (whether dryer balls are used: no in the left boxplot and yes in the right boxplot). The side-by-side display makes it easy to compare the medians, first and third quartiles, minimum, maximum, and possible outliers (there is no outlier in this data set) for the two levels of this categorical variable. Here, it appears as if these summary statistics are very similar and differ only by one or two minutes. It is important that the side-by-side boxplots are drawn using the same scale (which is done here) to allow for a fast visual comparison.

Figure 3 (center) shows a scatterplot of two quantitative variables. The different levels of a categorical variable (representing "yes" and "no") are shown via different colors and plotting symbols ("red circles" and "blue squares," respectively). Overall, an exponential increase can be seen in this graph. This scatterplot suggests it makes no major difference whether dryer balls are used or not used. A fourth categorical variable could be added via different plotting symbols (e.g., x, +, or o), while a fourth quantitative variable could be added via different sizes of the plotting symbols (where the area of the plotting symbol, and not its diameter, should be proportional to the numeric value).

Figure 3 (bottom) shows a line graph (time series plot) for three levels of a categorical variable for 27 time points. The data have been standardized to 1 at day 0. Typically, we standardize to 1 or 100%. While there is hardly any difference among the hardness of bread for the different factor levels during the first three days, the three factor levels result in different increases, starting with Day 4. While two of these increases could be described as quadratic or exponential (for the red and blue factor levels), the third factor level (green) shows a fairly linear increase of hardness.

Other Noteworthy Graphs

There are a few other graphs that should be mentioned here. When we want to describe the relationship between an explanatory variable (shown on the horizontal axis) and a response variable (shown on the vertical axis), we frequently add some smooth line to a scatterplot. This can be a least-squares regression line, as shown in the scatterplot in Figure 4 (top); a polynomial, exponential, or logarithmic fit to the data; or even a smoothed curve if no other relationship between the two variables can be found easily.

Histograms, such as the one shown in Figure 4 (center), are an effective way to summarize one quantitative variable. It is important that not too many and not too few classes are used. Several recommendations exist in the literature as to how many classes should be used. Moreover, it is recommended that all classes are equally wide. If this is not possible, it is necessary to translate the data into a density scale, where the vertical axis represents the count (or percentage) per unit on the horizontal axis.

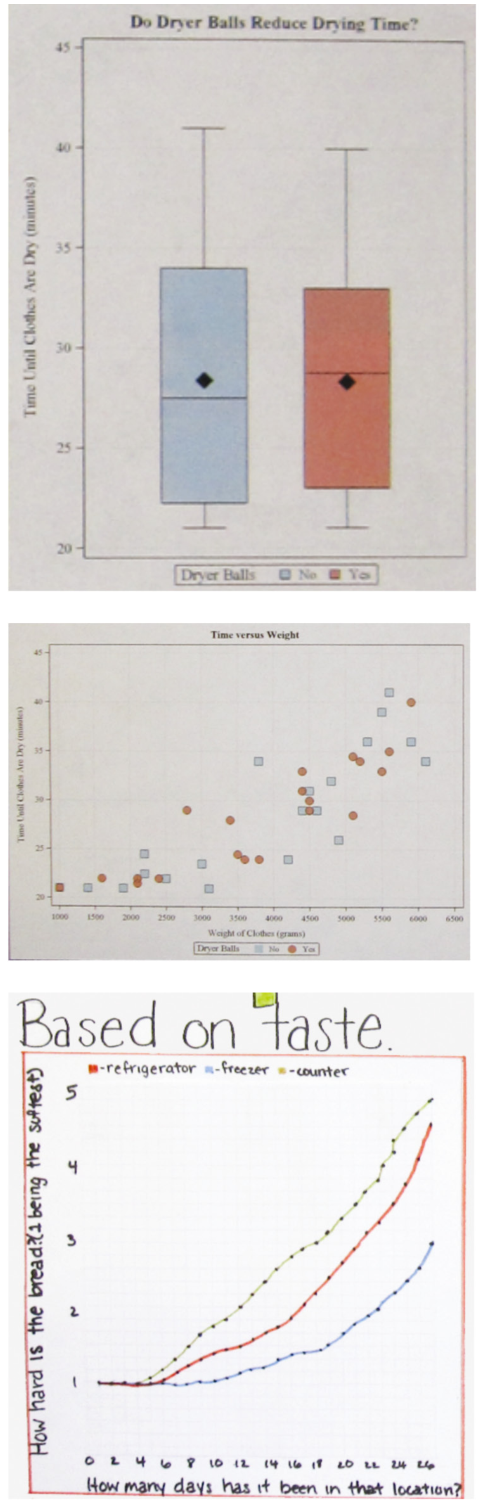


Figure 3: Excellent graphs for comparison: (top) side-by-side boxplots, (center) scatterplot with an additional categorical variable, and (bottom) line graph (time series plot) of three levels of a categorical variable

This histogram shows some interesting bimodal distribution properties, with peaks at 0–3 months and 12–15 months the reader may not have expected. That the population in this study consists only of high-school students at the junior and senior levels may serve as a possible explanation for this bimodality.

Whenever data have a geographic (spatial) component, such as the regions of the Atlantic Ocean shown to the west and north of Norway in Figure 4 (bottom), a map is of high importance. Figure 4 (bottom) shows a choropleth map, where the statistical information for a geographic region (such as countries or administrative districts) is shown via color (or shading). Choropleth maps are frequently used to show income in different sub-regions, election outcomes, or climatic data (such as temperatures and precipitation). It is much easier to comprehend geographic (spatial) information when shown on a map, compared to only a textual description of the geographic locations.

Conclusion

A variety of excellent graphs were used in the winning posters of the 2013 Statistics Poster Competition in all age groups. However, with respect to pie charts and bar charts, some versions of these graphs are preferred to other versions of the same graphs.

For More Information: Books and Articles

Note that graphs also are called charts, plots, or (statistical) graphics in the literature and on the web. Therefore, you may have to use one of these alternative terms (instead of graphs) when searching for further information.

Figures from *www.amstat.org/educationposterprojects/2013posters.cfm* (see also *Amstat News*, August 2013, pp. 26–33, *http://magazine.amstat.org/wp-content/uploads/2013an/August2013.pdf*). Photos are courtesy of the American Statistical Association.

Robbins, Naomi B. 2013. *Creating more effective graphs*. Chart House: Ramsey, NJ. (reprinted from Wiley 2005)

Su, Yu-Sung. 2008. It's easy to produce chartjunk using Microsoft Excel 2007, but hard to make good graphs. *Computational Statistics and Data Analysis* 52:4594–4601.

Tufte, Edward R. 2001. *The visual display of quantitative information*, 2nd edition. Graphics Press: Cheshire, CT. (first edition 1983)

Wainer, Howard. 2000. *Visual revelations: Graphical tales of fate and deception from Napoleon Bonaparte to Ross Perot*. Psychology Press: London, UK. (reprinted from Copernicus 1997)

For More Information: Blogs

Camoes, Jorge. ExcelCharts, *www.excelcharts.com/blog*.

Peltier, Jon. Peltier Tech Blog – Peltier Tech Excel Charts and Programming, *www.peltiertech.com/WordPress*.

Robbins, Naomi B. Effective Graphs, *www.forbes.com/sites/naomirobbins*.

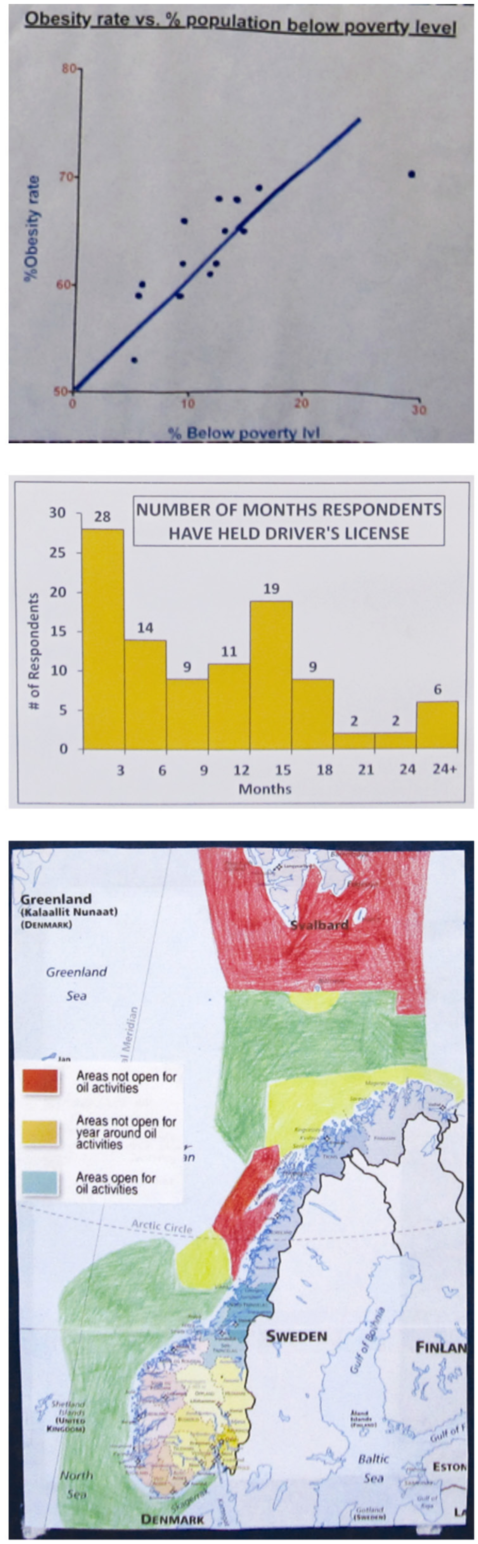


Figure 4: (top) Scatterplot with regression line, (center) histogram, and (bottom) choropleth map

My Experience Explaining Biostatistics to Elementary School Children

David Kerr

I volunteered to discuss my work as a biostatistician involved in clinical trials with my daughter's elementary school classroom and other 3rd, 4th, and 5th graders (8–11 years old) at her school. My intent was to give the students a sense of the drug approval process and clinical trials, as well as a general sense of statistical thinking. By using an example of a real clinical trial (with hands-on participation), I believe the students were able to understand what I and other biostatisticians do for a living. Hopefully, I inspired a few of them to explore their talents in math and statistics.

The following shows the flow of discussion with the students. Materials prepared in advance were packages of red and blue candy, enough pennies to distribute to every student, and slips of paper as found in Appendix A. The intent was to have discourse with the students and keep them engaged. I've included the text I used with the students, along with some additional comments about how the students responded and some of my reasoning for the wording, as well as how you might use this idea.

Photo: ThinkStock

Dialogue

"Hi there. I'm Mr. David, but some of you know me as Allison's dad. So my first question to you is: Have any of you ever gotten sick?"

> Lots of students' hands went up, and they told about things that have happened to them.

"Has anyone here ever taken any medicine when they were sick?"

> I intentionally never used the word "drug" during the presentation, only "medicine" and "medication." I didn't want to confuse the issue between 'bad' drugs and 'good' drugs. Lots of hands went up, with the students saying they'd taken aspirin or Benadryl, etc.

"Where did you get the medicine from?"

> Lots of hands went up as they gave me names of stores.

"Right, some medicines you can pick up right off the shelf in the Safeway or CVS or Walgreens. Some medicines you have to go to the back room of these stores—the pharmacy—and have a note from your doctor. And some medicines would only be given to you at a hospital. But all of these medicines have been approved by the government—the Food and Drug Administration (the FDA)—that the medicine should work and should be safe for you to take. But these medicines can be powerful, so you still only want to take them if your doctor or parent says so and you follow the instructions on the label.

"100 years ago, anyone could mix things together in a bathtub and sell it as a medicine to people. We don't let that happen anymore. Now, companies have to prove to the government that the medicine will work and is safe before they can start selling it at Walgreens or give it to you at a hospital. The government makes companies run experiments to prove that the medicine works and is safe. My company helps design these experiments and collect the data and analyze them. I study the data and work with doctors to help decide whether the new medicine really is safe and whether it really does help people get better. It makes me feel good. If the new medicine works, we can help it get to the stores quickly so that sick people can take it and get better. And if we discover the new medicine doesn't work, we make sure that no one else takes it.

"So suppose a scientist has an idea for a medicine that will help with a disease or illness. They try it in their laboratory and it seems like it should work. Then, they try it in some rats that have the disease and it seems like the rats get better. But, they have to try it in some people before they can start selling at a pharmacy or on a grocery store shelf. That is called a clinical trial. We test the new medicine in a small number of sick people to see if it works before we let everyone start buying it. We give some people the new medicine, and we give the rest something else. If the people on the new medicine do a lot better than the people taking the other thing, then we feel pretty sure the new medicine really helped.

"There are three important things for our clinical trials. First, we randomize. Some people will get the new medicine, some will get something else. If you help out by participating in this clinical trial, you will be randomly chosen and some people will get one treatment and some people will get another. This helps make sure the people in the two groups are similar to each other at the beginning. Second, we try to hide it so you don't know which treatment you're on. The people in our clinical trials don't know if they're on the new medicine or the other thing. Third, we compare things. One common thing we do is have the other group take a placebo. Does anyone here know what a placebo is? That could just be a pill with sugar inside. The placebo is like a fake medicine. So, some people are randomly given the new medicine and some are randomly given the placebo, and the people in the trial don't know which they are on. Then, at the end, we compare to see which group did better, the new medicine group or the placebo group. Hopefully, the people on the new medicine do better than the people on the placebo.

"Okay, I've got bad news.… You're all sick! You've got a cold. Feel free to cough (into your arm!), and let's hear some sniffles."

> The children had a good (noisy!) time pretending to be sick.

"But, I have good news! Our scientist back at the lab has discovered a new medicine that she thinks will help. She thinks it will shorten the length of time that you are miserable with your cold. But, most importantly, what should the name of this new medicine be?"

> I received lots of input from the children here. We went with "Maxisipe" (the classroom teacher's name was Ms. Sipe).

"So all of you sick children will participate in this clinical trial. Some of you will get Maxisipe and some of you will get the placebo. You'll take your one pill and we'll see how long it takes for you to get better. Then we'll see if the people who took Maxisipe get better quicker than those of you who took the placebo.

"First off, remember that we randomize people in our clinical trials. I'm going to hand out an indestructible metallic randomization device that is yours to keep. Can you guess what that is?"

> I handed out pennies. In my case, the children had a great time with them—flipping them, losing them on the floor, etc.—until the teacher finally was able to calm them down.

"All right, normally we would keep the randomization behind the scenes, but for this clinical trial, you will have a bit of information. So, everyone give your penny one final flip, which will determine which group you are in."

> The students flipped their pennies one final time.

"Before we go on, how many of the 24 people in our trial do you think flipped their coin as heads? Probably about half—around 12 people—but it could be a bit more or a bit less. It's random, but it would still be surprising to see everyone get heads or everyone get tails.

"Okay, everyone who had a heads, raise your hand."

> I went around the room and handed out RED candy to those who flipped up a heads. Yes, I let them eat it.

"Okay, who had a tails on their coin?"

> I handed out BLUE candy to those who flipped up a tails.

"So, some of you are on the RED treatment, and the rest of you are on the BLUE treatment. You know that much. But I'm not going to tell you yet which color matches up with Maxisipe and which matches up with placebo.

"Remember, medicine is not candy. Only take medicine if your parent or doctor says to. What would you do if you found some pills or something that looked like candy while you visited your grandma's house, or on the way to school? Would you eat it, or give it to your dog or little sister? No! Give it to your parent or teacher to safely throw away.

"Now, let's see how long it took everyone to get better."

> I handed out slips of paper as provided in Appendix A. These should be printed on a color printer (with the red and blue text). The slips of paper tell how many days it took each participant in our clinical trial to recover from the case of the cold. The numbers are rigged so that the RED group gets better with an average of 3 days, while the BLUE group gets better with an average of 4 days, but there is some variability in the data.

"Let's go around the room and have everyone in the RED group tell us how many days it took to get better. And now let's have everyone in the BLUE group tell us."

As the numbers were read, I entered the results into two ad hoc dot plots (shown to the class using the overhead projector) like the following:

RED

Count	Number of Days it Takes to Get Better						
6							
5		X					
4		X	X				
3		X	X				
2	X	X	X				
1	X	X	X	X	X	X	
	1 Day	2 Days	3 Days	4 Days	5 Days	6 Days	7 Days

BLUE

Count	Number of Days it Takes to Get Better						
6							
5							
4				X			
3		X	X	X			
2		X	X	X		X	
1	X	X	X	X	X	X	
	1 Day	2 Days	3 Days	4 Days	5 Days	6 Days	7 Days

"Which group seems to have recovered from their colds quicker?"

Let kids say RED.

"Why do you say RED? Both groups have the same minimum and maximum time to get better.… Tell me more about why it seems to you that RED is better at reducing the time to get better.

"I know you've been studying some good math and statistics ideas. Let's find the median and mean and mode for the number of days it takes for each of the groups to get better."

Note: Depending on the skill level of the kids, they might be able to do the calculations; otherwise, you will have to do them yourself. You also may need to explain why these are calculated. For example, you might say this is a way of summarizing a large number of data values into a single number. It's a lot easier to compare two groups when you can reduce the data to comparing one single number to another single number. The mode is the most common result, and the median is the number that has half the people above it and half the people below it. But, in this situation, the most useful for statisticians is the average, or mean.

"So it seems like the RED group is curing the cold faster, and maybe not just by coincidence. I'm sure you want to know, so I'm going to tell you now that the RED group was indeed our new medicine, Maxisipe. But, it's not enough to just have RED have a lower average. That could happen by chance. We want to make sure that it really is because of the new medicine and not just that the people who got RED were luckier than the BLUE group. Statisticians like me have mathematical tools that can say whether this difference was just by chance, or whether it's probably because of the new medicine.

"But, did anyone have anything else on their slip of paper?"

Some of the slips had a funny side effect listed in addition to saying how many days it took to get better. Examples included "All of your hair fell out," "Your brown eyes turned blue," "You got a rash all over your body," and "You lost your sense of smell."

"Let's go around the room and see how many people in the RED group had something bad happen to them and how many in the BLUE group had something bad happen. We call these side effects, or sometimes adverse events."

The slips were rigged so that about 30% of the RED group had an adverse event noted (about 5 students) and only 12% of the BLUE group had an adverse event (about 2 students). Write down the count and compute the percentage in each of the two groups that had an adverse event.

"Again, there could just be bad luck that results in more side effects in one group. But, if we start to see a trend with a lot more side effects of certain types in the people who took the new medicine, it would certainly get our attention as we review the data and try to decide whether to allow the medicine to be approved for people to use. Sometimes, we know that there will be side effects. Medicines to help people with cancer can cause lots of bad things, like making your hair fall out or throwing up a lot. But, we're okay with that if the medicine cures the cancer. But, if the medicine is intended to help with a cold, we wouldn't want there to be any bad side effects from our new medicine, since people will recover from their cold okay eventually, even without the medicine.

"Does anyone else have any examples of things like this happening when they've taken a medication?"

Lots of hands were raised and examples provided, like grandma making a meal while asleep after taking a sleeping pill and someone having an allergic reaction to a medicine.

"So, let's do a survey. I told you that the RED group was the new medicine, Maxisipe. It helped shorten the average time you were coughing and sniffling by about 2 days. But there are also a few more bad things that happened to people in that group. So, if you were sick, how many of you would want to take this new medicine?"

Students discussed whether to approve the new medicine.

"Should the FDA approve this new medicine to help reduce the time you are sick with the cold? That's the kind of question that statisticians and other experts in clinical trials have to try and answer. And now you have a better idea of all the hard work that goes on to make sure that the medicines you get will help you get better without hurting you.

"So, my official job is called a statistician. That's someone who helps decide what data to collect and then looks at data and tries to understand what the data are saying. It's kind of like being a detective. I always liked math and puzzles, and eventually went to college and, after a while, realized that statistics was a great field for me.

"Statistics is a great career because you get to work in a lot of different areas. My area is to see if new medicines work, but my friends use data from other areas. One of my friends uses data to help predict how many fish there are in the sea so we know how many fish the fishermen should be allowed to catch. And another one of my friends works at a bank and looks at data and creates computer programs that can figure out if someone has stolen your credit card number and is trying to buy something with it by seeing if the purchase is really different than other things you've bought before. The job pays pretty well also. So, if you like math and trying to find the truth, keep studying and perhaps you can be a statistician too."

Conclusion

It seemed like the children were engaged. In particular, they loved to tell about themselves—the medicines they've taken and the side effects they've experienced. They had a good sense of the problems faced here and were receptive to the notion of taking a smaller sample to help determine the risk/benefit profile before opening up the medicine to a larger group. I encourage others to volunteer in schools to help advocate for statistical literacy and, in particular, the methodology for clinical trials.

Notes for Teachers

My hope is you'll find the above material explanatory enough for you to present the ideas to your class, even if you have had no previous experience in clinical trials. The math involved is computing the mean, median, and mode, which can be found in most math textbooks. For further information about clinical trials, a brief overview is available at *http://clinicaltrials.gov/ct2/about-studies/learn*. Another option is to contact the American Statistical Association (*info@amstat.org*) to find a statistician who will come in and help facilitate this or describe their specific career.

APPENDIX A:

RED (cut into slips and fold over)

You got better in	4	days.	
You got better in	2	days.	
You got better in	1	days.	(But your hair fell out)
You got better in	3	days.	(But you couldn’t fall asleep at night)
You got better in	2	days.	
You got better in	2	days.	(But your brown eyes turned blue)
You got better in	3	days.	
You got better in	6	days.	
You got better in	2	days.	
You got better in	4	days.	(But you got a rash all over your body)
You got better in	3	days.	
You got better in	5	days.	
You got better in	2	days.	(But you lost your sense of smell)
You got better in	2	days.	
You got better in	1	days.	
You got better in	2	days.	
You got better in	3	days.	
You got better in	5	days.	
You got better in	4	days.	
You got better in	3	days.	
You got better in	2	days.	
You got better in	2	days.	
You got better in	3	days.	
You got better in	1	days.	
You got better in	5	days.	

BLUE (cut into slips and fold over)

You got better in 4 days.

You got better in 3 days.

You got better in 1 days. (But you couldn't stop hiccupping)

You got better in 5 days.

You got better in 3 days.

You got better in 5 days.

You got better in 7 days.

You got better in 4 days.

You got better in 2 days. (But you got a rash all over)

You got better in 6 days.

You got better in 5 days.

You got better in 5 days.

You got better in 2 days.

You got better in 4 days.

You got better in 3 days.

You got better in 4 days.

You got better in 3 days.

You got better in 5 days.

You got better in 3 days.

You got better in 4 days.

You got better in 5 days.

You got better in 2 days.

You got better in 3 days.

You got better in 3 days.

You got better in 5 days.

APPENDIX B:

Bar chart template:

RED

Count	Number of Days It Takes to Get Better						
6							
5		X					
4		X	X				
3		X	X				
2	X	X	X				
1	X	X	X	X	X	X	
	1 Day	2 Days	3 Days	4 Days	5 Days	6 Days	7 Days

Mean = ____ Median = ____ Mode = ____

Adverse event (side effect) rate is ____ out of _____ (____%)

BLUE

Count	Number of Days It Takes to Get Better						
6							
5		X					
4		X	X				
3		X	X				
2	X	X	X				
1	X	X	X	X	X	X	
	1 Day	2 Days	3 Days	4 Days	5 Days	6 Days	7 Days

Mean = ____ Median = ____ Mode = ____

Adverse event (side effect) rate is ____ out of _____ (____%)

Lessons from the LOCUS Assessments (Part 1): Comparing Groups

Douglas Whitaker and Tim Jacobbe

Since its publication, the *Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report: A Pre-K–12 Curriculum Framework* has been influential in the field of statistics education. The developmental levels—A, B, and C—through which students are hypothesized to progress provide convenient touchstones for curriculum and lesson design.

Despite the impressive contributions to statistics education in terms of instructional recommendations and the aforementioned developmental progression, the GAISE framework says less about what types of assessments are recommended or should be considered as a model. The NSF-funded Levels of Conceptual Understanding in Statistics (LOCUS) project focused on developing statistical assessments in the spirit of the GAISE framework. These assessments emphasize conceptual (rather than procedural) understanding and can be used to classify students as having understanding at level A, B, or C.

The assessments—which will be available in August 2014—consist of four forms: a pre- and post-test targeting the A and B levels and a pre- and post-test targeting the B and C levels. The A/B assessment was designed for students in grades 6–9, and the B/C assessment was designed for students in grades 10–12. Two versions of these are available—one with 23 multiple choice items and 5 free response items and another with 30 multiple choice items only. The items from which these assessments were constructed were piloted in spring 2013 with a total of 2,075 students for the A/B assessment and 1,249 students for the B/C assessment. (Although every item was not piloted with every student, each item was piloted with several hundred students.) While the pilot administration sample was large and included students of many backgrounds and ability levels, it was not selected to be a representative sample of students in the United States. We do report some overall performance indicators, but these are included to paint a more complete picture of the students and item and should not be over-interpreted.

Student work can be a valuable resource for teachers. The size and scope of the LOCUS pilot assessments yielded considerable variation in student responses. While there were some 'textbook' correct answers, students also were able to demonstrate correct statistical reasoning in imaginative ways. Incorrect answers often illustrated specific misunderstandings and, if identified as such, can suggest areas for more attention.

The LOCUS free response item being examined here is shown in Figure 1. This item addresses the following Common Core State Standards (CCSS):

- 6.SP.2 "Develop understanding of statistical variability."
- 6.SP.5 "Summarize and describe distributions."
- 7.SP.3 "Draw informal comparative inferences about two populations."
- S-ID.1-3 "Summarize, represent, and interpret data on single count or measurement variable."
- S-IC.3 "Making inferences and justifying conclusions."

The "Analyze" and "Interpret" components of the GAISE framework at Level B also are addressed by this item.

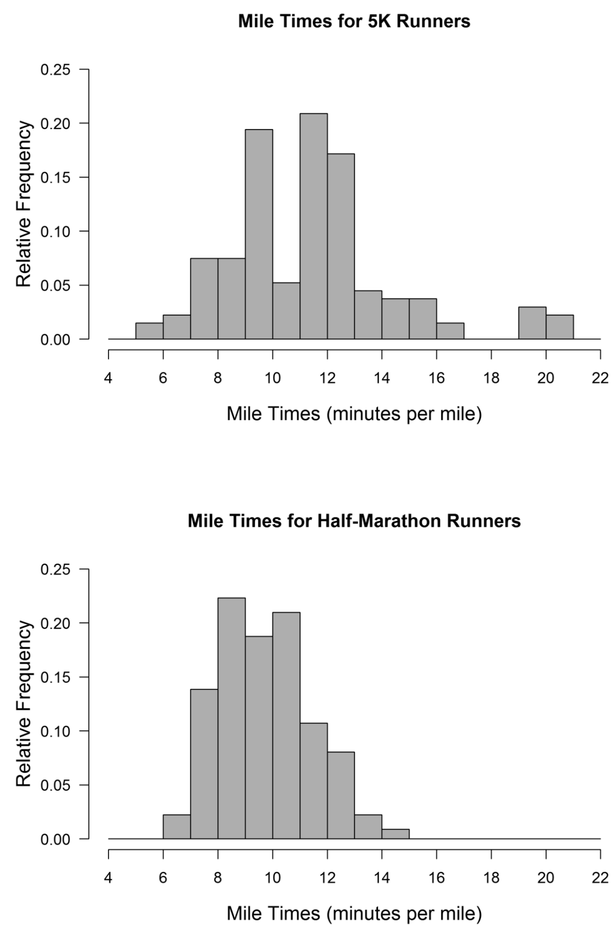


Figure 1: The city of Gainesville hosted two races last year on New Year's Day. Individual runners chose to run either a 5K (3.1 miles) or a half-marathon (13.1 miles). One hundred thirty four people ran in the 5K, and 224 people ran the half-marathon. The mile time, which is the average amount of time it takes a runner to run a mile, was calculated for each runner by dividing the time it took the runner to finish the race by the length of the race. The histograms show the distributions of mile times (in minutes per mile) for the runners in the two races.

Student Responses

This free response item was piloted with a total of 618 students in grades 9–12. Free response items were scored out of 4 points, with a 4 indicating a "complete" response (allowing for mistakes such as minor computational errors not indicative of a misunderstanding). For each item, a small team of graders established a rubric and conducted initial item scoring aloud as a group. Once every grader felt comfortable with applying the rubric, scoring continued individually. Any discrepancies or questions were brought to the group's attention. The distribution of students' scores for the item is given in Table 1: 4.8% earned a 4, 10.8% earned a 3, 33.8% earned a 2, 23.1% earned a 1, and the remaining 19.6% earned either a 0 or did not provide a response (these do not sum to 100% due to rounding).

Score	Percent of Students
4	4.8%
3	10.8%
2	33.8%
1	23.1%
0 or no response	19.6%

Table 1. The distribution of student scores for the item. (These do not sum to 100% due to rounding.)

Question related to free response item in Figure 1:

> *Jaron predicted that the mile times of runners in the 5K race would be more consistent than the mile times of runners in the half-marathon. Do these data support Jaron's statement? Explain why or why not.*

The first piece of this question related to Figure 1 is asking about how the mile times of 5K runners compare to the mile times of half-marathon runners—as a group. To answer the question correctly, students need to (1) recognize this is a question comparing groups and not individuals and (2) be able to correctly interpret histograms. Many students incorrectly focused on the heights of the bars representing relative frequency and concluded that variability in bar heights implies inconsistent mile times, as in this student response: "No, because there are more spikes in the graph for the 5K, and less in the graph for the half marathon."

Other researchers have discussed this misconception. Linda Cooper and Felice Shore, in a *Journal of Statistics Education* article, found that nearly 50% of students in their sample of 186 undergraduates judged variability in histograms by focusing on the heights of bars and attribute this to the visual similarity of histograms to bar charts and time-plots that use bars. The term "spread" is used often as a synonym for variability; while this word may seem accessible, students may be inclined to focus on the evenness of bars.

To score full marks on this question, a student must attend to some measure of the horizontal (rather than vertical) variability in the data. The measure of variability used need not be sophisticated. One student attempted to use the range with reasonable results: "No, the data in fact supports the opposite: that the half-marathon times are more consistent than the 5K times. The half-marathon times as you can see range from 6–15 minutes, whereas the range of 5K times is 5–21 minutes and 30 seconds, a much larger range (9 to 16.5)." This student also used the range, but made an implicit comparison: "No, the data doesn't support Jaron. The mile times for the 5K runners have a larger range and a higher standard deviation." Even though the student did not explicitly mention the half-marathon runners, language such as "larger" and "higher" indicate a comparison is being made.

Question related to free response item in Figure 1:

> *Sierra predicted that, on average, the mile time for runners of the half-marathon would be greater than the mile time for runners of the 5K race. Do these data support Sierra's statement? Explain why or why not.*

This question asks the students to directly compare the mile times for the two groups. Many students made appropriate arguments based on the mean or median: "No, the mean time for the half marathon is approximately about 9–10 minutes, where the 5K mean time is approximately about 11–12 minutes. And the 5K has some much higher times, which will increase the mean."

Of the students who did not compare the centers of the two groups, this response exemplifies a common misconception: "Yes, because the times stayed consistent during the 8–11 mile times." This student is focusing on the proportion of runners in the half-marathon group whose times were between 8 and 11 minutes. This approach attends to only part of the data in one group, does not make an appropriate comparison with the 5K runners group, and seems to be based on reasoning using the mode rather than a measure of center for quantitative data such as the mean or median.

Question related to free response item in Figure 1:

> *Recall that individual runners chose to run only one of the two races. Based on these data, is it reasonable to conclude that the mile time of a person would be less when that person runs a half-marathon than when he or she runs a 5K? Explain why or why not.*

The core component of this question is that the way runners were assigned to run either the 5K or half-marathon matters and has real implications for the conclusions that can be drawn from the data. As this student responded, "No, different races attract people of different abilities. Since these people ran their races voluntarily (they weren't randomly selected), nothing can be concluded." Although the student confused the terminology *random selection* with *random assignment*, they demonstrated a clear understanding that the runners chose either the 5K or half-marathon and that there could be a valid reason why a runner's mile time would not be less in a half-marathon than a 5K. Students were imaginative and many potentially valid reasons were given, all of which were scored as correct.

Some students correctly indicated concluding that mile times would be less when running a half-marathon than a 5K is inappropriate based on the given data, but their reasoning was incorrect: "No, because to do that you need to have an individual run both races to compare differences in time." While a matched-pairs design would work to answer the question at hand, it is not strictly necessary. This student's response ignores the lack of random assignment that could allow such a conclusion to be reached in a properly designed study.

In many cases, students can benefit from drawing on their experience and prior knowledge for answering questions. Other students, however, can be led astray by this if they rely on it too heavily: "Yes, if a person chose the half marathon over the 5K you can assume they are in better shape and are better runners than someone who would have chose to run less in the 5K." While such an assumption about the runners' abilities influencing their choice may turn out to be true in some cases, it is not appropriate to reach a conclusion on the basis of an untested assumption.

Discussion

This item broadly targets the "Analyze" and "Interpret" components of the GAISE framework and specifically targets several CCSS standards (6.SP.2, 6.SP.5, 7.SP.3, S-ID.1-3, S-IC.3). For students needing help with the content covered in this item, there are many resources available, including lesson plans on STEW (*www.amstat.org/education/stew*). Several lesson plans on STEW address the key aspect of comparing two groups using graphical displays (e.g., "How Long Is 30 Seconds?" and "Colors Challenge!"). Box plots are the graphical display used in these lesson plans, but histograms could be included and address the same CCSS standards (6.SP.4 and S-ID.1).

It is worth noting that, while this item was viewed favorably by the LOCUS development team and provided many interesting student responses, there were two complications that led to it not being included on the final version of the assessment. First, the histogram for the 5K runners has both a larger range (and other traditional measures of variability) and is 'bumpier' than the histogram for the half-marathon runners. Thus, when a student response included the half-marathon runners are more consistent with weak or confused justification, it was difficult to determine if the student was demonstrating a misconception about interpreting a histogram or simply not providing adequate justification.

Second, some students seemed to have difficulty with the concept of *mile time*—the measure used to compare the runners of races with different lengths in this problem. This presented confusion as to what was meant by a "greater" mile time: Does this mean larger in magnitude (slower) or better (faster)? The purpose of this question was not to test understanding of the mile time concept, but to test conceptual understanding of statistics. As such, it was replaced on the final forms of the assessment by questions that did not have these complications. The hundreds of student responses as written, though, still proved valuable.

Further Reading

Cooper, L. L., and F. S. Shore. 2008. Students' misconceptions in interpreting center and variability of data represented via histograms and stem-and-leaf plots. *Journal of Statistics Education* 16(2). Retrieved from *www.amstat.org/publications/jse/v16n2/cooper.html*

Franklin, C., G. Kader, D. Mewborn, J. Moreno, R. Peck, M. Perry, and R. Scheaffer. 2007. *Guidelines for assessment and instruction in statistics education (GAISE) report: A pre-k–12 curriculum framework*. American Statistical Association: Alexandria, VA. Retrieved from *www.amstat.org/education/gaise*

Kaplan, J. J., D. G. Fisher, and N. T. Rogness. 2009. Lexical ambiguity in statistics: What do students know about the words association, average, confidence, random, and spread? *Journal of Statistics Education* 17(3):n3.

National Governors Association Center for Best Practices and Council of Chief State School Officers. 2010. *Common core state standards for mathematics*. National Governors Association Center for Best Practices and Council of Chief State School Officers: Washington, DC.

Shaughnessy, J. M. 2006. Student work and student thinking: An invaluable source for teaching and research. In *Proceedings of the Seventh International Conference on Teaching Statistics*. Retrieved from *http://iase-web.org/documents/papers/icots7/PL6_SHAU.pdf*

The Gumball Machine

Written by:
Alexander White, M. Alejandra Sorto, and Rini Oktavia
Texas State University - San Marcos
sorto@txstate.edu

Overview of Lesson

This lesson gives students the opportunity to explore and discuss the variation that occurs in sampling. Students are asked to 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. Then students are asked several questions about the probabilities of certain events that could occur.

This lesson has four goals: (1) to understand that events associated with small probability can and do occur; (2) to realize that the distribution of data from small samples often does not reflect the parent distribution; (3) to take variation into account when describing possible sampling outcomes; and (4) to recognize an effect of sample size on sampling variation. In order to lead students to focus more on variation than point estimates, students are asked to imagine that 5 students select 10 gumballs each and to predict the number of blue gumballs the students get (see Figure 1).

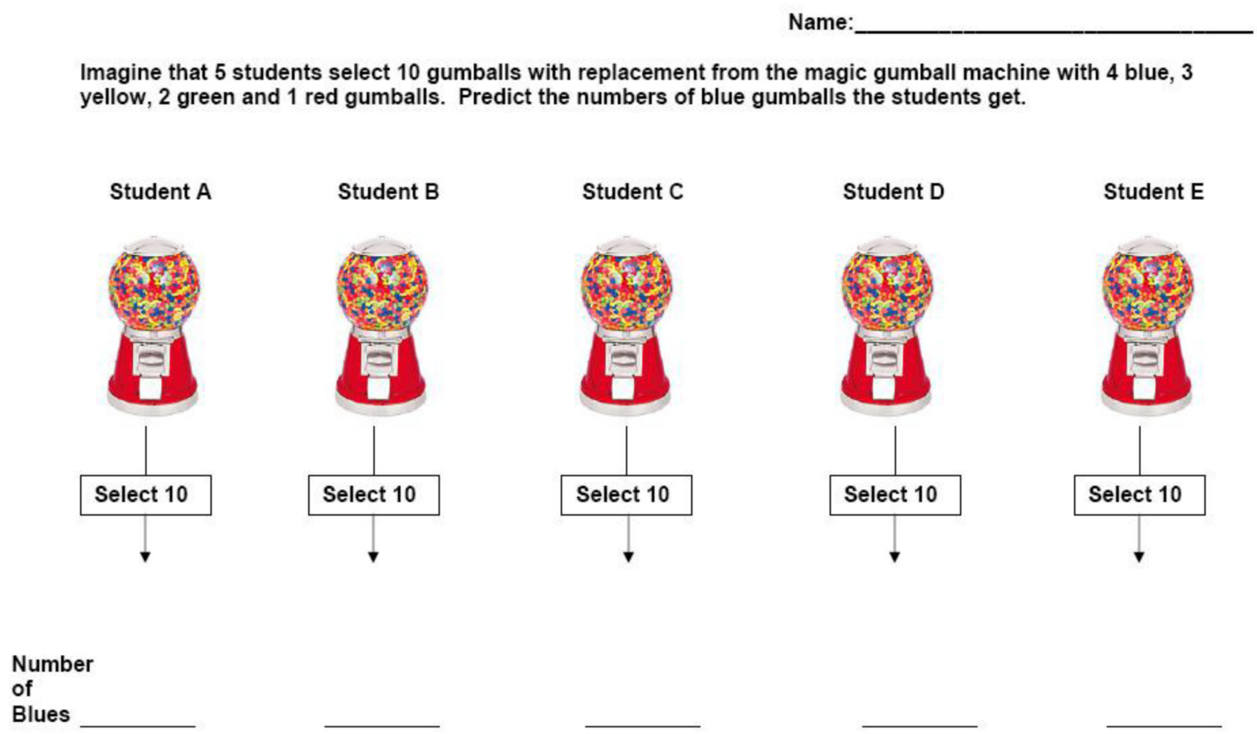


Figure 1. Visualizing the gumball machine.

To see the effect of sample size, students are asked to find and plot the experimental probability for both samples of $n = 10$ gumballs and $n = 20$ gumballs; and to help visualize variability across many samples, students are 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. After a discussion about the prediction of the outcomes, students do the experiment. After students record their outcomes, they compare with their previous predictions. Based on previous use of the lesson conducted, we find that most students are surprised to see that their predictions were, in past students' own words, "way off".

GAISE Components

This investigation follows the four components of statistical problem solving put forth in the *Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report*. The four components are: formulate a question, design and implement a plan to collect data, analyze the data by measures and graphs, and interpret the results in the context of the original question. This is a GAISE Level A activity.

Common Core State Standards for Mathematical Practice

1. Make sense of problems and persevere in solving them.
2. Reason abstractly and quantitatively.
3. Construct viable arguments and critique the reasoning of others.

Common Core State Standard Grade Level Content (Grades 3 through 6)

3. MD. - 5. MD. Represent and interpret data.

6. SP. 1. Recognize a statistical question as one that anticipates variability in the data related to the question and accounts for it in the answers.

6. SP. 4. Display numerical data in plots on a number line, including dot plots, histograms, and box plots.

6. SP. 5. Summarize numerical data sets in relation to their context by: reporting the number of observations.

NCTM Principles and Standards for School Mathematics

Data Analysis and Probability Standard for Grades 3-5

Formulate questions that can be addressed with data and collect, organize, and display relevant data to answer them:

- design investigations to address a question and consider how data-collection methods affect the nature of the data set;
- collect data using observations, surveys, and experiments;
- represent data using tables and graphs such as line plots, bar graphs, and line graphs;
- recognize the differences in representing categorical and numerical data.

Select and use appropriate statistical methods to analyze data:

- describe parts of the data and the set of data as a whole to determine what the data show.

Develop and evaluate inferences and predictions that are based on data:

- use proportionality and a basic understanding of probability to make and test conjectures about the results of experiments and simulations.

Prerequisites

Students should know how to place fractions (in particular tenths and twentieths) on the number line.

Learning Targets

Students will be able to graphically display data and interpret data.

Time Required

1 class period.

Materials Required

Color sticky notes, a cup, color chips, pencil and paper, and a copy of the Activity Sheet (pages 7 through 9).

Instructional Lesson Plan

The GAISE Statistical Problem-Solving Procedure

I. Formulate Question(s)

Begin the investigation by asking students to 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. Then ask students to imagine that 5 students select 10 gumballs each with replacement (i.e. select one gumball, observe its color, replace the gumball in the machine, repeat this process 10 times). To encourage students to develop their abilities in formulating questions, ask students to write some questions that they would be interested in investigating about the number of red, green, yellow, and blue gumballs that the 5 students will get.

The class generates a list of questions, which the teacher records on the board. Some possible questions might be:

- What is the list of possible numbers of blue gumballs that the five students get?
- What number of blue gumballs showed up the most?
- What color of gumballs is chosen the most?

Have the students think about how the number of red, green, yellow, and blue gumballs that each student gets will vary. Then ask students to answer the following questions:

1. If you got one gumball from the machine, which color do you think would come out?
2. If we sampled 10 gumballs from the magic gumball machine, how many gumballs of each color will come out?
3. Imagine that 5 students pick 10 gumballs each, predict the number of blue gumballs student A will get? How about student B? Student C, D and E?

II. Design and Implement a Plan to Collect the Data

Give students a copy of the "Gumball Machine" activity worksheet.

First Task. (Problem 1) Ask one of the students to read out loud the first question ("If you got one gumball from the machine, which color do you think would come out?"). Demonstrate the gumball machine and the gumballs using a cup and color chips. Then, ask students to individually answer the first question and explain their reasoning in writing on their worksheet. After a few minutes, asks students to share their responses. It is more likely that students will say that a blue gumball will come out because the machine has more blue gumballs than any other color. Even though "blue" is the best prediction for a single outcome, it is important that students are aware of the complement of this event. Therefore, as a follow-up question, ask students: "What are the chances of getting a color different than blue?" and "Are the chances of getting a color different than blue the same as getting blue?". Ask students to record their answers.

Second Task. (Problem 2) Ask students to answer 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, act 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 may show a large variety of reasoning. For example, a few students may say things such as "you don't know" or "you never know" because "it will be restored each time" or "because you take one out and you are putting them back in the machine". They may show an "equiprobability bias" (Lecoutre 1992) that is the belief that since we are selecting randomly: anything can happen and everything has the same chance of occurring. Other students may say "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". Other students may give a different distribution of colors such as "3 blue, 4 yellow, 1 green, 2 red," or "6 blue, 2 yellow, 1 green, 1 red" or outcomes that contain less than 10 total gumballs. Other students may predict "4 blue, 3 yellow, 2 green, 1 red" and their explanations are based mainly on the original distribution of the colors in the gumball machine.

Third Task. (Problems 3 - 5) This task focuses only on the number of blue gumballs. The question addressed is "Imagine that 5 students select 10 gumballs each, what is your prediction of the number of blue gumballs the students get?" Phrasing the prediction question in this fashion will give the students the opportunity to explore and discuss the variability we would expect in sampling. Students' predictions for the numbers of blue gumballs will vary. Most students may provide a sequence of numbers varying from 1 to 9. Using statistical theory as a guide, the predictions can be characterized as "low", "high", "wide" and "narrow". Low predictions are a sequence of numbers with means less than 3 and high predictions are those with means greater than or equal to 6. A "narrow" prediction is a sequence of repeated numbers (e.g. "6, 6, 6, 6, 6") that suggest no variability in the sampling results. A sequence is "wide" if its range is greater than or equal to 6 (Shaughnessy, Ciancetta, and Canada 2004).

Give each pair of students a cup with 10 colored chips and ask them to simulate the gumball machine. Have the students follow the directions given on the activity sheet. Make sure the students sample with replacement and shake the cup between each selection. Students will have outcomes for 10 trials. After the students record their outcomes, compare the outcomes to their previous predictions.

III. Analyze the Data

Ask students to repeat the experiment one more time (Problem 6). Then have students combine their data from both experiments (Problem 7). So, students will have outcomes for samples of size 10 and size 20. In problems 8 and 10 on the worksheet, students will compute the fractions of blue, yellow, green, and red that they get for both 10 and 20 trials. This allows for a discussion of experimental probability and the possibility to compare the results from the two different sample sizes by placing the values on a number line (Problems 9 and 11). To show the variability of the experimental probabilities, have each student place a blue dot sticker representing their observed results for n = 10 trials on a pre-made large number line (see Figure 2).

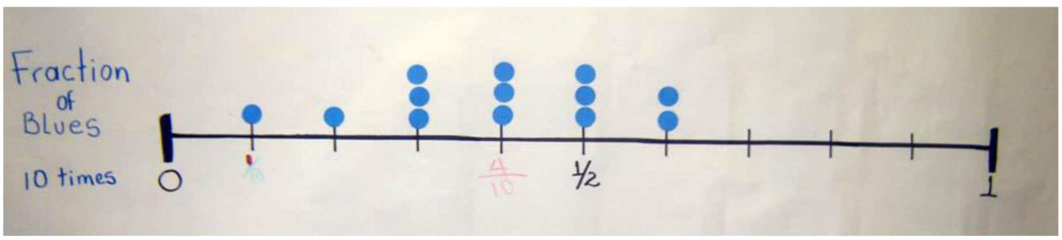


Figure 2. Dot plot of observed experimental probabilities

Use a separate number line to plot the results for n = 20. Compare the results from the two dot plots.

IV. Interpret the Results

Use the dot plots as the spring board for a discussion about what occurred when selecting the samples of gumballs. For example, you may ask the students "Did anyone get 4 blues?" and "Where do you see this on the graph?" Compare the results to their predictions about the number of blues for the 5 imaginary students. Discuss why it is not likely to get 4 every time, but it is not likely to get 0 or 10. The discussion should not focus only on the center of the resulting distribution but the spread and extreme values as well. This will set the basis for more formal reasoning about variation down the road. Have the students compare the results for n = 10 and n = 20. By comparing the results, students should see that with a larger sample the experimental probabilities are less spread out.

Assessment

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.

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.

Answer:

1. (b)
2. (a)

Possible Extensions (to GAISE Level B)

1. After the dot plots are constructed, ask students to describe the shape, center and spread of the distribution of the number of blue gumballs and to examine the distribution for clusters, gaps, and potential outliers.
2. Ask students to construct a boxplot to display the data distribution and to determine if there are any outliers using the 1.5xIQR rule.
3. Ask students to calculate and discuss measures of center (mean, median).
4. Ask students to discuss the spread of the distribution using common measures of spread: range, interquartile range, mean absolute deviation.

References

1. Lecoutre, M. P. (1992), "Cognitive models and problem spaces in 'purely random' situations," *Educational Studies in Mathematics*, 23, 557- 568.
2. Shaughnessy, J.M., Ciancetta, M., and Canada, D. (2004), "Types of Students Reasoning on Sampling Tasks," *Proceedings of the 28th Conference of the International Group for the Psychology of Mathematics Education*, 2004 Vol. 4, pp. 177 – 184.
3. Sorto, M. A. & White, A. (2008). The Gumball Machine: Linking Research and Practice about the concept of variability. *Journal of Statistics Education*. 16(2). *www.amstat.org/publications/jse*

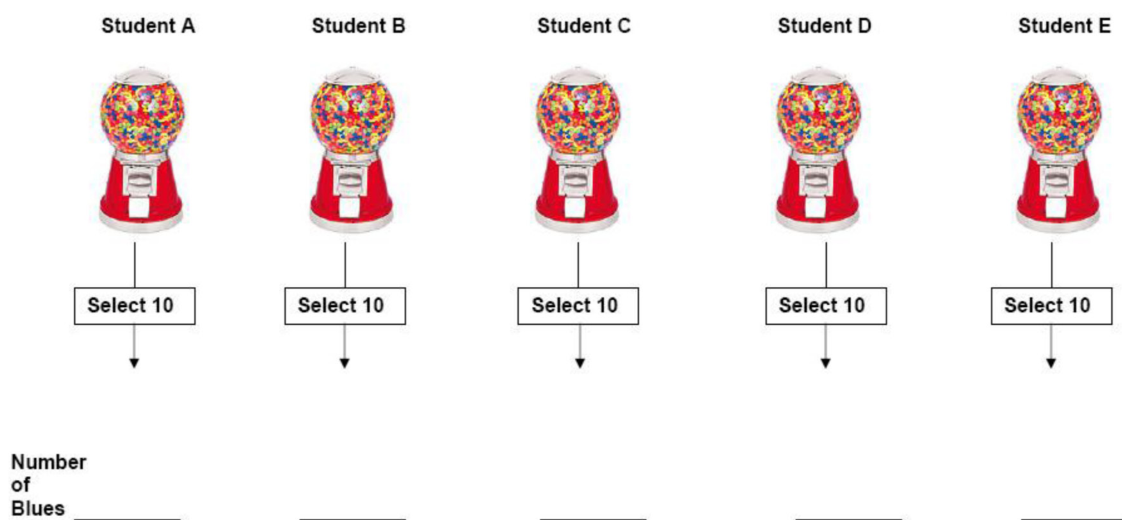
The Gumball Machine Activity Sheet

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?
Blue ___ Red _____ Green ______ Yellow _______
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.



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 (out of 10)				

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

6. Repeat the experiment one more time and record your data below.

Color of Chip	Blue	Yellow	Green	Red
Tally				
Number (out of 10)				

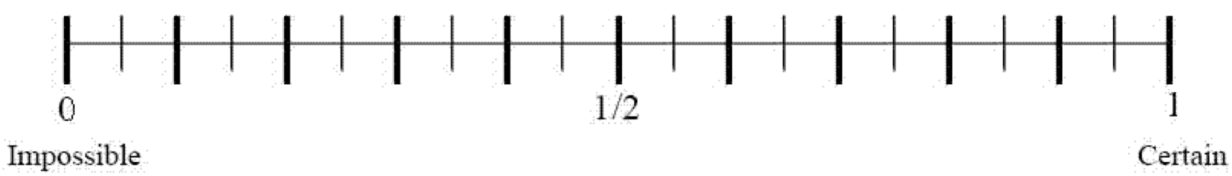
7. Combine the results for all 20 trials below.

Color of Chip	Blue	Yellow	Green	Red
Number (out of 20)				

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				

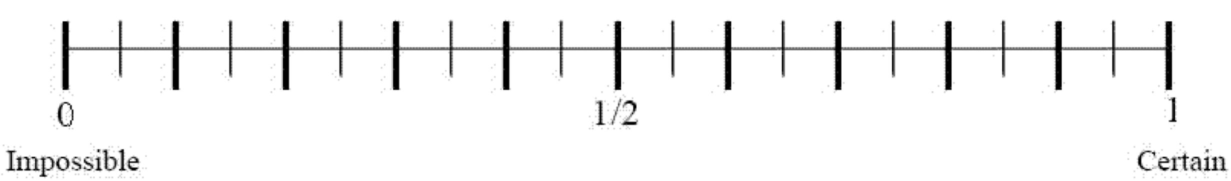
9. Use B, Y, G and R to represent the four colors. On the scale below, show the experimental probability of drawing each color you got in Task 8.



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 10.



ANNOUNCEMENTS

Registration Now Open for 2nd Biennial Electronic Conference on Teaching Statistics (eCOTS)

The registration system for the 2nd Biennial Electronic Conference on Teaching Statistics (eCOTS) is now open. Please go to *www.causeweb.org/ecots* to learn about our exciting program and register for eCOTS. The conference will take place entirely online, through CAUSEweb, from May 19–23. In addition to many sessions and posters, there are six workshops you can sign up for. There also will be keynote talks from Christine Franklin (University of Georgia) and Conrad Wolfram (Wolfram Group). The three themes of eCOTS are "Teaching with Big Data," "The Impact of the Common Core," and "Bridging the Disciplines." Registration for eCOTS is $25.

Meeting Within a Meeting (MWM) Statistics Workshop for Middle- and High-School Mathematics and Science Teachers – Boston, Massachusetts – August 5–6

MWM will take place in conjunction with the Joint Statistical Meetings this summer in Boston, Massachusetts. The workshop is meant to strengthen K–12 mathematics and science teachers' understanding of statistics and provide them with hands-on activities aligned with the Common Core State Standards that they can use in their own classrooms. The cost of the workshop is $50. Online registration will be available at *www.amstat.org/education/mwm*.

Beyond AP Statistics (BAPS) Workshop – Boston, Massachusetts – August 6

The ASA/NCTM Joint Committee is pleased to sponsor a Beyond AP Statistics workshop at the annual Joint Statistical Meetings. Organized by Roxy Peck, the BAPS workshop is offered for experienced AP Statistics teachers and consists of enrichment material just beyond the basic AP syllabus. The cost of the workshop is $50. Online registration will be available at *www.amstat.org/education/baps*.

2014 Statistics Project Competition

Introduce grades 7–12 students to statistics through the annual project competition (*www.amstat.org/education/posterprojects*), directed by the ASA/NCTM Joint Committee on Curriculum in Statistics and Probability. The competition offers opportunities for students to formulate questions and collect, analyze, and draw conclusions from data. Winners will be recognized with plaques, cash prizes, certificates, and calculators, and their names will be published in *Amstat News*. Projects are due June 1.

Judges Sought for Statistics Project Competition

The ASA/NCTM Joint Committee on Curriculum in Statistics and Probability is seeking judges for the 2014 Statistics Project Competition (*www.amstat.org/education/posterprojects*). Judging takes place via email during the summer and requires about four hours of your time. If interested, please email Daren Starnes, head judge, at *dstarnes@lawrenceville.org*.

Episode 5 of STATS+STORIES Is Available

Episode 5 ("Health, Privacy, and Confidentiality") of S+S is now available at *www.statsandstories.net*. This is a conversation with Paul Scanlon, a survey methodologist and research social scientist in the Questionnaire Design Research Laboratory at the Centers for Disease Control and Prevention's National Center for Health Statistics. You can subscribe to the Stats+Stories podcast on iTunes.

New Online Community for ASA K–12 Teacher Members

A new online community for ASA K–12 Teacher Members will allow participation in online discussions and sharing resources with other members. More information is available at *http://community.amstat.org/participate/helpfaqs*. Not yet an ASA K–12 Teacher Member? You can start your free three-month trial at *www.amstat.org/membership/K12teachers*.

ASA Writing Statistics Education of Teachers Report

In light of the Common Core State Standards, the Conference Board of the Mathematical Sciences (CBMS) released *The Mathematical Education of Teachers II* (*MET2*), which focuses on the mathematics and statistics preparation of K–12 teachers. The ASA review of *MET2* was well received by CBMS, which encouraged the ASA to expand recommendations to a white paper. The ASA board recently funded this project—led by Christine Franklin and Tim Jacobbe, both members of the ASA/NCTM Joint Committee—to create a companion report on the statistics education of teachers. For more information, see *http://magazine.amstat.org/blog/2014/03/01/education-of-teachers*.

New The World of Statistics Website and Resources

The free international statistics education resources created during the 2013 International Year of Statistics are now available and ongoing through The World of Statistics website. Teachers everywhere can access a wealth of statistics instruction tools and resources from around the world at *www.worldofstatistics.org*.

ANNOUNCEMENTS

Census at School Program Reaches 20,000 Students

The ASA's U.S. Census at School program (*www.amstat.org/censusatschool*) is a free, international classroom project that engages students in grades 4–12 in statistical problemsolving. The students complete an online survey, analyze their class census results, and compare their class with random samples of students in the United States and other participating countries. The project began in the United Kingdom in 2000 and now includes Australia, Canada, New Zealand, South Africa, Ireland, South Korea, and Japan. The ASA is seeking champions to further expand the U.S. Census at School program nationally. For more information about how you can get involved, read the article at *http://magazine.amstat.org/blog/2012/02/01/censusatschool-2* or email ASA Director of Education Rebecca Nichols at *rebecca@amstat.org*.

Free Statistics Education Webinars

The ASA offers free webinars about K–12 statistics education topics at *www.amstat.org/education/webinars*. This series was developed as part of the follow-up activities for the Meeting Within a Meeting Statistics Workshop. The Consortium for the Advancement of Undergraduate Statistics Education also offers free webinars about undergraduate statistics education topics at *www.causeweb.org*.

Project-SET

Project-SET is an NSF-funded project to develop curricular materials that enhance the ability of high-school teachers to foster students' statistical learning regarding sampling variability and regression. All material is geared toward helping high-school teachers implement the Common Core State Standards for statistics and are closely aligned with the learning goals outlined in the *Guidelines for Assessment and Instruction in Statistics Education: A Pre-K–12 Curriculum Framework (GAISE)* report. For more information, visit *http://project-set.com*.

Provide Feedback for Updating Curriculum Guidelines for Undergraduate Statistics Programs

The ASA is updating the curriculum guidelines for undergraduate statistics programs. The working group—consisting of representatives from academia, industry, and government—welcomes your input at *www.amstat.org/education/curriculumguidelines.cfm*.

UPCOMING CONFERENCES

Electronic Conference on Teaching Statistics (eCOTS)
www.causeweb.org/ecots
The next eCOTS will be held online May 19–23.

Joint Statistical Meetings (JSM)
www.amstat.org/meetings/jsm/2014
August 2–7, Boston, Massachusetts

The Meeting Within a Meeting Statistics Workshop for Math and Science Teachers (*www.amstat.org/education/mwm*) and **Beyond AP Statistics** Workshop (*www.amstat.org/education/baps*) will be held at JSM 2014.

The 9th International Conference on Teaching Statistics (ICOTS 9)
http://icots.net/9
July 13–18, Flagstaff, Arizona

STEW Lesson Plan

For this and other free, peer-reviewed lessons, please visit *www.amstat.org/education/stew.*

Enable Students to Explore Real Data on TuvaLabs

TuvaLabs empowers students to think critically about data, ask meaningful questions, and communicate their observations and findings. Teachers can conduct inquiry-driven activities, lessons, and projects related to various subjects including math, science, social studies, health, and others aligned to the Common Core Content and Practice Standards. With TuvaLabs (*www.tuvalabs.com*), teachers can equip their students with the knowledge, values, and skills to become active members of their own communities and global citizens of the world. TuvaLabs is free, and always will be for teachers and students.

MWM
Meeting Within a Meeting
a statistics workshop for K-12 teachers

Mathematics and Science Teachers

(*www.amstat.org/education/mwm*)

Sponsored by the American Statistical Association (ASA) 2014 Joint Statistical Meetings (JSM)*

Based on the Common Core State Standards for Mathematics (*corestandards.org*) and *Guidelines for Assessment and Instruction in Statistics Education (GAISE): A Pre-K–12 Curriculum Framework* (*www.amstat.org/education/gaise*)

Dates: Tuesday, August 5 and Wednesday, August 6, 2014, 8:00 a.m. to 4:00 p.m.

Place: Boston, Massachusetts, Boston Convention and Exhibition Center or a nearby conference hotel (workshop meeting room location to be announced)

Audience: Middle and High School Mathematics and Science Teachers. Multiple mathematics/science teachers from the same school are especially encouraged to attend. Note: Experienced AP statistics teachers should register for the Beyond AP Statistics (BAPS) workshop. See *www.amstat.org/education/baps* for more information.

Objectives: Enhance understanding and teaching of statistics within the mathematics/science curriculum through conceptual understanding, active learning, real-world data applications, and appropriate technology

Content: Teachers will explore problems that require them to formulate questions, collect, organize, analyze, and draw conclusions from data and apply basic concepts of probability. The MWM program will include examining what students can be expected to do at the most basic level of understanding and what can be expected of them as their skills develop and their experience broadens. Content is consistent with Common Core standards, *GAISE* recommendations, and *NCTM Principles and Standards for School Mathematics*.

Presenters: *GAISE Report* authors and prominent statistics educators

Format: Middle school and high school statistics sessions

One-day pass to attend activities at JSM (statistics education sessions, poster sessions, JSM exhibit hall)

Activity-based sessions, including lesson plan development

Provided: Refreshments

Complimentary one-day pass to attend the Joint Statistical Meetings

Handouts

Certificate of participation from the ASA certifying professional development hours

Optional graduate credit available

Cost: The course fee for the two days is $50. Please note: Course attendees do not need to register for the Joint Statistical Meetings to participate in this workshop.

Follow up: Follow-up activities and webinars (*www.amstat.org/education/k12webinars*)

Network with statisticians and teachers to organize learning communities

Registration: More information and online registration is available at *www.amstat.org/education/mwm.* Space is limited. If interested in attending, please register as soon as possible.

Contact: Rebecca Nichols, *rebecca@amstat.org*; (703) 684-1221, Ext. 1877

*The Joint Statistical Meetings are the largest annual gathering of statisticians, where thousands from around the world meet to share advances in statistical knowledge. The JSM activities include statistics education sessions, posters sessions, and the exhibit hall.

BAPS

BEYOND AP STATISTICS

Wednesday, August 6, 2014
8:00 a.m. - 4:30 p.m.
Boston, Massachusetts

A WORKSHOP FOR EXPERIENCED TEACHERS

Sponsor: ASA-NCTM Joint Committee on Curriculum in Statistics and Probability

The ASA/NCTM Joint Committee is pleased to sponsor a Beyond AP Statistics (BAPS) workshop at the annual Joint Statistical Meetings* in Boston, Massachusetts on August 6, 2014. Organized by Roxy Peck, the BAPS workshop is offered for AP statistics teachers and consists of enrichment material just beyond the basic AP syllabus. The course is divided into four sessions led by noted statisticians.

PRESENTERS:

Allan Rossman and Beth Chance, Cal Poly - Inference for Paired Data

Tom Short, John Carroll University, Logistic Regression

Robin Lock, St. Lawrence University - What Do We Do When Assumptions Are Not Met?

James Cochran, Louisiana Tech University - Engaging Students in Statistics

COST:

The course fee for the full day is $50. Please note: Course attendees do not need to register for the Joint Statistical Meetings (JSM)* in order to participate in this workshop, although there is discount JSM registration for K-12 teachers available at *www.amstat.org/meetings/jsm/2014*.

LOCATION:

Boston, Massachusetts, Boston Convention and Exhibition Center or a nearby conference hotel
(workshop meeting room location to be announced)

PROVIDED:

Refreshments (lunch on your own)
Handouts
Pass to attend the Exhibit Hall at the Joint Statistical Meetings
Certificate of participation from the American Statistical Association (ASA) certifying professional development hours
Optional graduate credit available

REGISTRATION:

More information and online registration is available at *www.amstat.org/education/baps*. Registrations will be accepted until the course fills, but should arrive no later than July 15, 2014. Space is limited. If interested in attending, please register as soon as possible.

QUESTIONS:

Contact Rebecca Nichols at *rebecca@amstat.org* or call (703) 684-1221 ext. 1877

*The Joint Statistical Meetings are the largest annual gathering of statisticians, where thousands from around the world meet to share advances in statistical knowledge. The JSM activities include statistics and statistics education sessions, posters sessions, and the exhibit hall.

STatistics Education Web

Lesson Plans Available on Statistics Education Web for K–12 Teachers

Statistics Education Web (STEW) is an online resource for peer-reviewed lesson plans for K–12 teachers. The lesson plans identify both the statistical concepts being developed and the age range appropriate for their use. The statistical concepts follow the recommendations of the *Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report: A Pre-K-12 Curriculum Framework*, *Common Core State Standards for Mathematics*, and *NCTM Principles and Standards for School Mathematics*. The website resource is organized around the four elements in the GAISE framework: formulate a statistical question, design and implement a plan to collect data, analyze the data by measures and graphs, and interpret the data in the context of the original question. Teachers can navigate the site by grade level and statistical topic.

Lesson Plans Wanted for Statistics Education Web

The editor of STEW is accepting submissions of lesson plans for an online bank of peer-reviewed lesson plans for K–12 teachers of mathematics and science. Lesson plans will showcase the use of statistical methods and ideas in science and mathematics based on the framework and levels in the *Guidelines for Assessment and Instruction in Statistics Education (GAISE)* and *Common Core State Standards*. Consider submitting several of your favorite lesson plans according to the STEW template to *steweditor@amstat.org*.

For more information, visit *www.amstat.org/education/stew*.

CENSUS at SCHOOL

FREE international classroom project to engage students in statistical problemsolving

Teach statistical concepts, statistical problemsolving, measurement, graphing, and data analysis using your students' own data and data from their peers in the United States and other countries.

Complete a brief online survey (classroom census)

13 questions common to international students, plus additional U.S. questions

15–20-minute computer session

Analyze your class results

Use teacher password to gain immediate access to class data.

Formulate questions of interest that can be answered with Census at School data, collect/select appropriate data, analyze the data—including appropriate graphs and numerical summaries for the corresponding variables of interest—interpret the results, and make appropriate conclusions in context relating to the original questions.

Compare your class census with samples from the United States and other countries

Download a random sample of Census at School data from United States students.

Download a random sample of Census at School data from international students (Australia, Canada, New Zealand, South Africa, and the United Kingdom).

International lesson plans are available, along with instructional webinars and other free resources.

www.amstat.org/censusatschool

For more information about how you can get involved, email Rebecca Nichols at *rebecca@amstat.org*.

Bridging the Gap

Between Common Core State Standards and Teaching Statistics

Twenty data analysis and probability investigations for
K–8 classrooms based on the four-step statistical process as defined by the
Guidelines for Assessment and Instruction in Statistics Education (GAISE)

www.amstat.org/education/btg