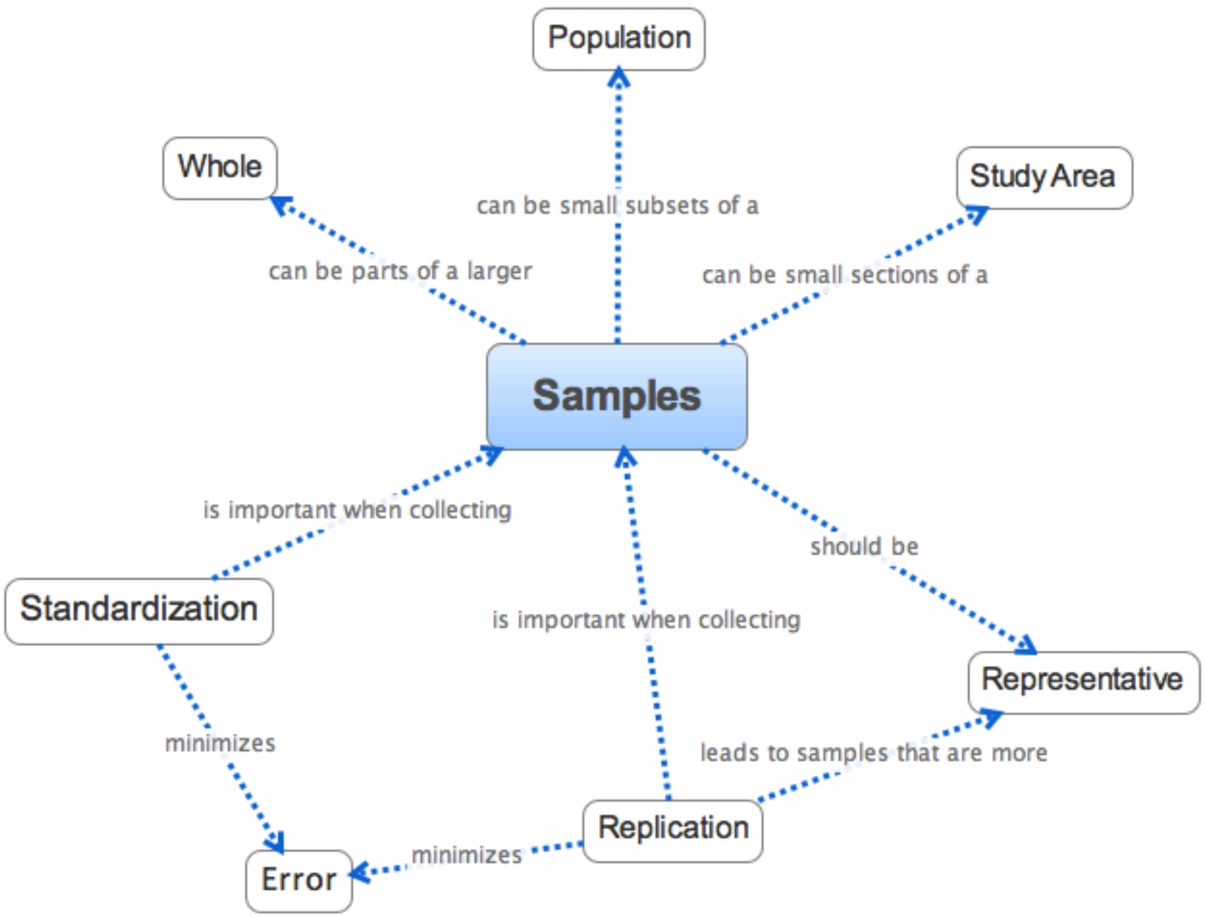
1.1 Sampling Design (SD)



T-SD Fig 1.1 Sampling Concept Map

Goals
Students will…

1. Sample a bag of small colored objects
2. Determine proportions of colors of small objects
3. Describe importance of standardization and replication in sampling

Ocean Literacy Principle(s) Addressed
Principle 7: *The ocean is largely unexplored*

- Understanding the ocean is more than a matter of curiosity. Exploration, inquiry and study are required to better understand ocean systems and processes. (OLP 7b)
- Use of mathematical models is now an essential part of ocean sciences. Models help us understand the complexity of the ocean and of its interaction with Earth's climate. They process observations and help describe the interactions among systems. (OLP 7e)

Background and Introduction
If using plain milk chocolate m&m's, the proportions of each color–the "entire population"–produced by the m&m's factory per the m&m's representative (February 2012) are as follows:

- Red: 13%
- Orange: 20%
- Yellow: 14%
- Blue: 24%
- Green: 16%
- Brown: 13%

The proportions produced differ for each type of m&m's candy (e.g. peanut, dark chocolate, pretzel), so check the m&m's website or call their customer service line if you are using a different product.

T-SD Table 1.1. Factory percentages of different types of m&m's candy (2012)

	Plain	Peanut	Dark Chocolate	Peanut Butter and Almond	Pretzel
Red	13	12	17	10	14.3
Orange	20	23	16	20	14.3
Yellow	14	15	17	20	14.3
Green	16	15	16	20	14.3
Blue	13	23	17	20	28.5
Brown	13	12	17	10	14.3

Chance factors involving the machines used by the manufacturer introduce random variation into the final color proportions of each bag produced. Some bags will have a distribution of colors that is close to the proportions produced by the factory, whereas others will be further away.

Upon completion of the activity, ask the class if they think the color proportions in their bag will be the same in other m&m's bags. The students sampled one bag many times, but the bag itself is only a sample of the "entire population" of m&m's produced at the factory. Thus, the class activity was really a sub-sampling of a sample (i.e. the bag) of the whole population (i.e. all of the m&m's produced at the factory).

Common Misconceptions

T-SD Table 1.2. Sampling misconceptions

Misconception	Explanation
Random Sampling A sample is a miniature version of the population.	A single sample will capture part of the population. If the population is not well-mixed, or if part of the population is very rare, one sample will not accurately represent the entire population. There is always variability between samples. This is why multiple samples are necessary.
Sample Size The larger the sample, the greater the accuracy.	The sample size needed for a study depends on the size of the population and the desired level of accuracy. For the same margin of error, a small population size needs a smaller sample size than a larger population.
Census vs. Sample Taking a census is better than sampling.	Taking a census is often time consuming, expensive, and unrealistic. A census can also miss critical parts of the population. A sampling study can be designed to capture both rare and common individuals, creating a more accurately representation of a population. The choice between census or sampling will depend on your study goals, the size of the population, and other resources.

From Saunders, M.N.K, Teaching how to sample in research: some things that might help.

Activity: Sampling Design

T-SD Table 1.3. Suggested activity progression, if adhering to closely to activity as written with no major modifications, assuming class periods of 40 minutes.

Day	Task
1	Introduction to activity
	Activity
2	Class Discussion

Optional Extension: Additional 20-30 min after final class discussion on activity

T-SD Table 1.4. Materials, if adhering closely to activity as written with no major modifications, assuming class of 32 students divided into groups of two (16 groups).

Materials	Quantity	Per	Class Total	Notes on Material Number or Material Modification
Small colored objects	Depends upon object	Class	Depends upon object	Objects could be colored candy, such as m&m's or jellybeans, colored plastic chips or paperclips, or other small, colored objects. If you use m&m's, refer to the section detailing m&m proportions in the background and introduction section. If putting together a sampling jar, for example with colored jellybeans, it is useful to know the quantities of the different colors. A jar of one pound of red jellybeans, half a pound of green jellybeans, and a quarter pound each of yellow and orange jellybeans works well for a class of 25 students. Thoroughly mix the jellybeans before the activity.
Bag or jar	1	Class	1	Unopened candy bag or clear bag or jar

Activity Inquiry Prompts
1. What is a sample?
2. Why do you predict that?
3. What will you do if your sample is too large or too small?
4. How would your results be different if your sampling method was different?

Procedure
NOTE Figures and tables in this section refer to figures and tables in EOFE unless noted.

Using a bag with known quantities is a very unusual and artificial sampling situation. However, because the quantities in the bag are known, this activity is a powerful way to demonstrate the concept of sampling and to show how sampling works.

Safety Warning(s): *If you are using candy, encourage the practice of food safety. Students can use gloves, utensils, and/or wash their hands.*

1. Write down all of the different colors that are in the bag. If you cannot see into the bag, use your prior knowledge to come up with a list of the colors of the objects in the bag.

Based on the content area you area teaching, you can refer to the different colors in discipline-specific terminology. For example, different colors can represent different temperature readings, elements in a soil sample, or species in a community.

If the bag your students are sampling contains colors that are not easily placed in separate categories, such as solid red jellybeans and red jellybeans with yellow spots, you can hold a discussion about what this might mean. For example, these jellybeans might be considered different isotopes of the same element, and thus be grouped together. Or the jellybeans could be considered different forms of the same species, perhaps juveniles and adults or males and females. When discussing categories, it does not matter what your students decide, as long as they reach a class consensus on how they categorize the objects. This is similar to the work of professional scientists who have to agree on classifications.

2. Make predictions about the colors of objects in the bag.
    a. Which color(s) do you think are most abundant?
    b. Which color(s) do you think are least abundant?

Write student predictions on a class table. Students can predict the most and least abundant color, they can rank the colors (1, 2, 3, etc.), or students can predict percentages of each color in the bag.

If using common objects in an opaque bag (e.g. m&m's) ask students to base their predictions on prior knowledge. If using unfamiliar objects, show students the clear jar of items so they have a visual baseline upon which to make their predictions.

3. Using your predictions, develop and record a hypothesis about the colors of objects in the bag.

4. Each person will be taking a sample of objects from the bag. Develop a standardized sampling scheme considering the following:
    a. How many objects should each person collect?

    The class can limit their sample by number (e.g. ten items) or volume (e.g. the capacity of a small scoop). For small objects, you may want to steer students to sampling by number or volume. For sampling by number, 10 is an easy sample size with which to do calculations.
    b. What will be used to collect the samples?

c. How will samples be removed from the bag?

For the purposes of this activity, as long as the procedure does not introduce additional bias and entire class agrees, the procedure is standardized. The specifics of the procedure do not matter. However, be careful of allowing students to choose a sample size that is so large they will "sample" the entire bag.

d. How will you ensure that the sampling is random?

For example, close eyes while sampling to avoid choosing particular colors, do not look while scooping, or remove the first items you touch.

e. What will you do if a person takes a sample that is larger or smaller than the size determined by the class?

Design could include taking another sample, closing eyes and randomly discarding extras, only count the ten closest items, etc.

5. Using the procedure developed by the class, sample the bag. Each member of the class should take one sample. Complete Table 2 with the colors of objects you sampled, the total number of each color, the total number of all the colors together, the fraction of each color compared to the total, and the percent of each color compared to the total.

Individual samples may not contain all colors in the larger population.

Table 2. Individual sample data table
Example table

	Color						**Total**
	Red	***Green***	***Yellow***				
Number	***6***	***3***	***1***				***10***
Fraction of Total	***6/10 or 0.6***	***3/10 or 0.3***	***1/10 or 0.1***				***10/10 or 1***
Percentage of Total (%)	***60%***	***30%***	***10%***				***100%***

6. Based on your individual sample,
   a. What can you infer about the colors in the bag?
   b. How do you think your sample may compare to other people's samples?

Student answers will vary based on how representative of the bag they think their sample is. Students may report that they can tell how many colors there are, which ones are the most common and which the least, and actual mathematical ratios of colors.

7. Record three samples from any three class members on a class data table. Table 3 is an example data table; you can modify the number of color columns and rows based on your class.

Record students' samples on a class data table everyone can see.

Student	Red (0.5)	Green (0.25)	Yellow (0.125)	Orange (0.125)	Total
1	*6*	*3*	*1*	*0*	*10*
2	*1*	*4*	*2*	*3*	*10*
3	*3*	*3*	*2*	*2*	*10*
Total	*10*	*10*	*5*	*5*	*30*
Average	*3.3*	*3.3*	*1.7*	*1.7*	*10*
Fraction	*3.3/10 or 0.33*	*3.3/10 or 0.33*	*1.7/10 or 0.17*	*1.7/10 or 0.17*	*10/10 or 1*
Percentage	*33%*	*33%*	*17%*	*17%*	*100%*

8. Based on the first three samples shared by the class,
   a. What can we infer about the colors in the bag if we stop our data collection at this point?
   b. What might we observe as we add more samples to the data table?

Based on this example, students may conclude that the proportions in the population are the same as those in the three samples, or they may conclude that more samples are needed. Stopping the data collection at this point illustrates the importance of taking many samples. The first few samples usually do not result in the same proportions as does the full data set.

9. Record the remaining samples on the class data table. For each color, determine the average, fraction, and percentage.

The following table is an example of data from eight student samples of a bag containing items of four different colors. The numbers in parentheses along the top row are the known proportions of each color. A fraction can be expressed as a decimal or as a simple fraction.

Percentages are calculated by multiplying the decimal proportion by 100. Percentages add up to 100%. You can use both percentages and proportions when comparing the class data to known color quantities.

Student	Red (0.5)	Green (0.25)	Yellow (0.125)	Orange (0.125)	Total
1	*6*	*3*	*1*	*0*	*10*
2	*1*	*4*	*2*	*3*	*10*
3	*3*	*3*	*2*	*2*	*10*
4	*5*	*4*	*0*	*1*	*10*
5	*4*	*1*	*3*	*2*	*10*
6	*8*	*1*	*0*	*1*	*10*
7	*5*	*2*	*2*	*1*	*10*
8	*4*	*3*	*1*	*2*	*10*
Total	*36*	*23*	*11*	*12*	*80*
Average	*4.5*	*2.9*	*1.4*	*1.5*	*10*
Fraction	*4.5/10 or .45*	*2.9/10 or .29*	*1.4/10 or .14*	*1.5/10 or .15*	*10/10 or 1*
Percentage	*45%*	*29%*	*14%*	*15%*	*100%*

As more samples are added to the table, students will be able to see that individual samples may not reflect the proportions of the bag very well. However, when all the samples are averaged together, they more accurately represent the proportion of colors in the bag. The more samples, the more accurate the data. This demonstrates the power of replication.

Activity Question Answers

1. What did you learn about the composition of the bag by collecting data on your individual sample?

Student answers will vary based on how representative of the bag they thought their sample was. Students may report they can tell how many colors there are, which ones are the most common and which the least, and actual mathematical ratios of colors.

2. What did you learn about the composition of the bag by compiling and averaging class data?

Sampling many times (replication) gives more information about the bag and leads to greater accuracy in estimating bag composition.

3. How did your individual data, the first three points of class data, and the entire set of class data compare? Explain.

This is a synthesis of the students' understanding of the completeness of the data as more data points are added. As more data points are added, accuracy of data increases.

4. Ask your teacher to share the known proportion of each color of object in the bag. How did the averaged class data proportion of colors compare to the known proportion of colors? Explain why you think this occurred.

Students may be amazed by how closely their class average samples match the bag's known proportions. Random sampling of the bag works because each sample gives a little more information about the colors in the bag.

5. What would happen to your sample data if you picked your favorite colors out of the bag when sampling?

If some students preferred the color red, and chose more red items, this would affect the results. This is the concept of bias, favoring some sample or part of the sample over another for some reason.

6. What situations in nature might cause individual samples to not reflect the larger area?

Answers will vary. For example, samples collected after a recent natural disturbance, like a storm or a fallen tree, or in an area heavily impacted by humans, such as around a picnic bench. In other cases, an environment may naturally vary greatly in one space, for example, an intertidal zone has areas that are exposed and submerged at different depths in a relatively small space. While samples might not reflect the larger area over time, they are important for giving specific information about a smaller area or event. Sometimes these situations are interesting and scientists will choose to focus on them. It depends on the research question.

7. What would happen to your sample data if there were just one or two objects of a particular color in the bag?

The probability of these items showing up in any student's sample would be very small. These items would represent something that is rare. Even after sampling the bag many times the rare color may not show up in anyone's sample. In nature, even using the best sampling practices, you may not capture all of the information about an area.

8. Read Practices of Science: Scientific Error.
    a. What are the sources of error in this color object sampling activity?
    b. How well do you think your sampling procedure controlled for bias? Explain.

Answers will vary. Error could include counting or calculation errors or losing sample, for example by dropping or eating candy. Bias was controlled by not choosing colors and worked to the extent that students followed the procedure.

9. How do you think this activity accurately represented sampling in a real environment, like the intertidal?

Answers will vary. There were multiple organisms. We didn't know how many of each type were in the bag. We had to take multiple samples to estimate population abundance.

10. How do you think this activity was different from sampling in a real environment, like the intertidal?

Answers will vary. We did not have to worry about the environment or the organisms actual safety. We were able to count the whole population (bag) and compare to our estimation (using out samples) to determine our accuracy (in real life it is impossible to know the absolute truth).

Optional Introduction

Goals

Students will…

1. Be introduced to the idea of sampling.
2. Compare the concepts of food sampling and scientific sampling.

T-CA Table 3.5. Materials, if adhering closely to demonstration as written with no major modifications

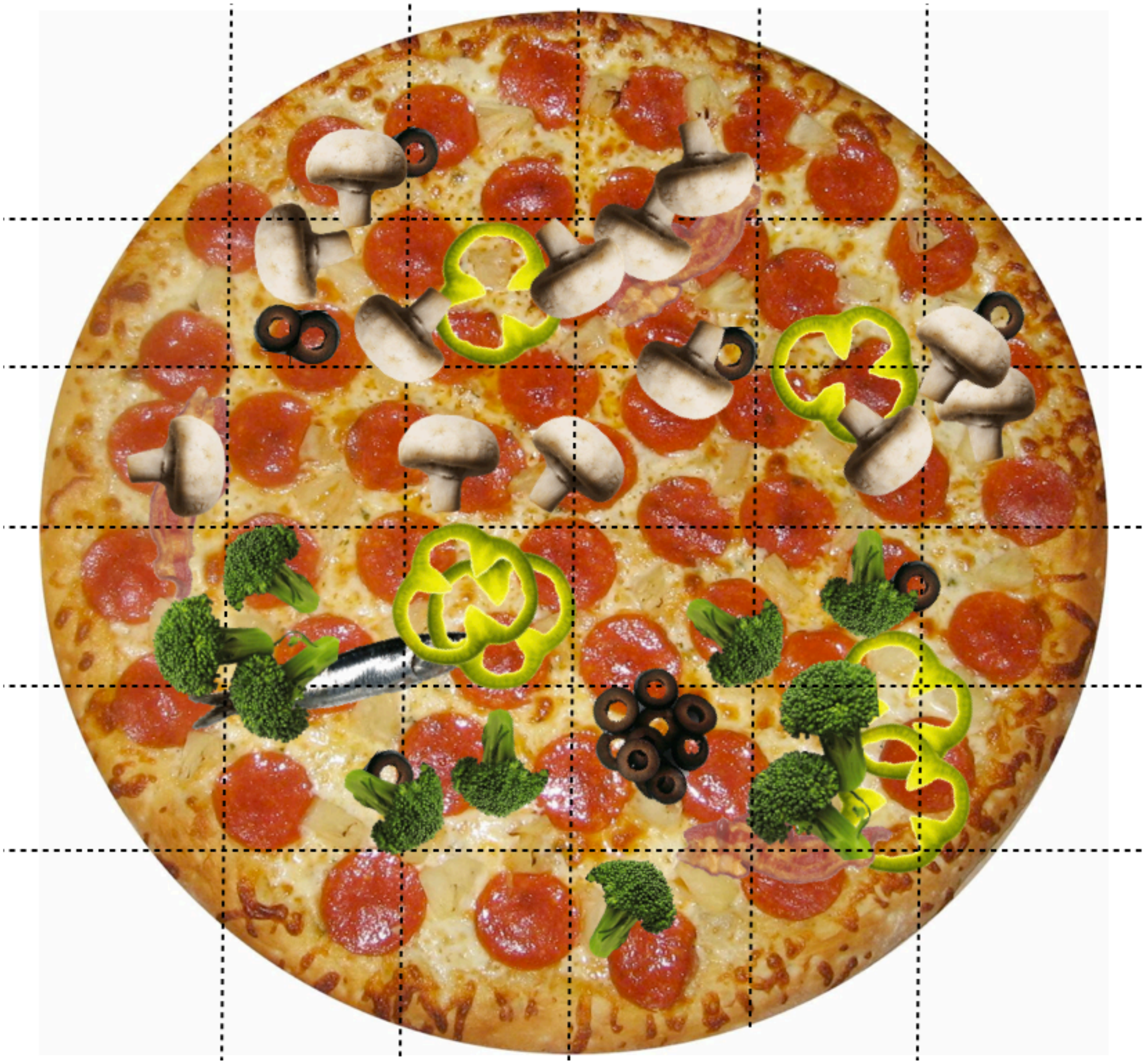
Materials	Quantity	Notes on Material Number or Material Modification
Pizza picture	1	Cut into enough small pieces for one per student; toppings should be distributed such that some are all over (e.g. cheese, pepperoni), some are only on half of the pizza (e.g. mushrooms, broccoli), and some are irregularly distributed (e.g. anchovy, olives)
Chef's Hat (optional)	1	See next section on how to make a chef's hat

Food Sampling Procedure

1. Draw or print out a picture of a pizza with many different toppings.
2. Cut the paper pizza into enough pieces for each student to get one piece, or sample.
3. Hand out a sample to each student.
4. Ask students to determine the toppings they think are on the pizza based on their individual samples.
5. Ask students to compare their samples to others in the group, then determine the toppings on the pizza based on the samples in their group.

Food Sampling Questions

1. What is a sample?
2. Did your individual piece of pizza tell you everything that was on the pizza?
3. What did you learn about the pizza as you compared your samples to other people's samples?



Sampling Introduction
Optional Introduction–Food Sampling: How to Make a Chef Hat

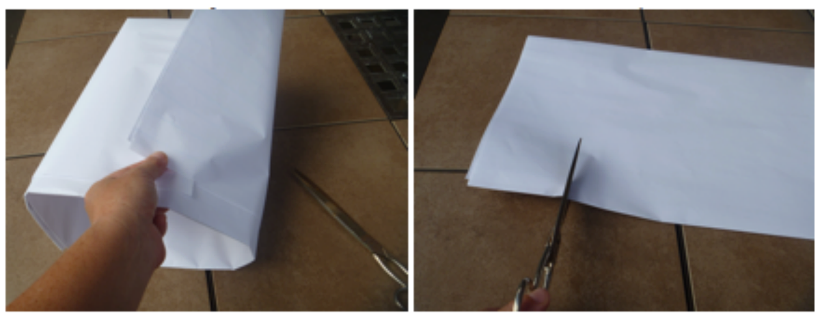
Materials:
- 2 sheets of chart paper or other large paper
- Clear tape
- Scissors

Procedure:
1. Fold the top of one piece of chart paper to the bottom of the paper so the unlined side is facing out. Fold the top of the paper to the bottom of the paper twice more. This will create a multi-layer paper band about 1/8th the height of the paper. This will be the brim of the hat.

2. Measure the circumference of your head with the brim band and tape to fit.


3. Fold a second piece of chart paper in thirds so the unlined side is facing out.

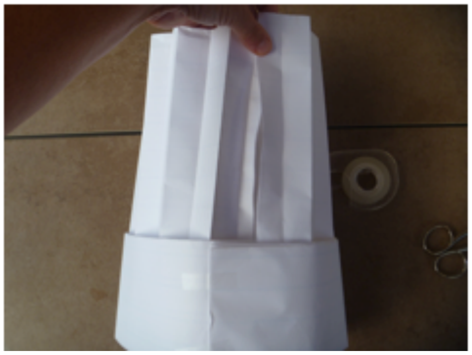
4. Trim the second piece of paper to match the circumference of the brim band. You may want to use the band to help measure the proper width.



5. Fold the now-trimmed second piece of paper accordion-style along the shorter side to create pleats.

6. Loosen the inner layer of the brim band and start to place the pleated paper between the inner and outer layers of the brim band.

7. Continue placing the pleated paper in the band all the way around the brim band and tape together the pleated ends when they connect.

8. Tape the pleated part of the hat to the brim band on the inside of the brim band.

9. Voila! You have a chef hat! Bon appetit!

Photos by A. Gundersen

Optional Extensions

Goals
Students will…

1. Learn more about bias
2. Replicate the m&m's sampling procedure

Bias Extensions (Activity Questions 3 and 6)
If students are having difficulty grasping the concept of bias, redo the exercise and introduce bias on purpose to see how it affects the data. Options:

- Have students select their favorite colors, which would throw off the calculated class proportions from the known quantities.
- Place colors into a jar in layers. Have students select ten items from the top. These will lead students to over-sampled the colors on top and conclude there are a larger proportion of them in the jar than there really is.

M&M Extensions
- Count all of the colors from the bag to see how well the class sample estimated the bag proportions.
- Do this activity again (replication) with different bags of the same type of M&Ms and average the color proportions. This will allow your class to better approximate the proportions produced in the factory.