SAMPLING AND PROBABILITY WORKSHOP 7: DISCOVERY

Agenda for Two-Hour Workshop

15 minutes *Workshop Facilitator/Site Leader*
Introduction

Hand out the materials for Workshop 7. Discuss the following questions:
- Why is the study of Sampling & Probability important in the middle school curriculum?
- How can we introduce Sampling & Probability so that it is relevant for our students?
- How is Sampling & Probability used to solve problems?

60 minutes *Whole Group*
View Workshop 7: Sampling & Probability — Discovery

While watching the program, consider the following focus questions:
- What is the difference between theoretical and experimental probability?
- What does it mean when the probability of an event happening is 0? 1?
- How many different ways can probabilities be stated or written?
- How many trials are necessary to be able to make an accurate prediction about an event?
- What are the two variables that are being graphed in Lesson 2: Graphing Probabilities?
- What type of graph (bar graph, line graph, etc.) is most appropriate for this display?
- What would be an appropriate scale for the x-axis? The y-axis?
- If you are graphing the number of trials versus the probability of an event, what would you expect to happen to the graph as you repeat more trials?
- What prerequisite skills do students need to be successful with these lessons?
- How can these lessons be used to reinforce basic skills?

40 minutes *Small Groups or with a Partner*
Read, Do and Discuss

Read Lesson 3: Predicting M&Ms.
Do Lesson 3.
Discuss findings, focusing on the following questions:
- What is a sample and how is it used?
- What makes a sample truly representative of the larger population?
- If you have a population of 100 items (like the M&M bags), what are some methods for determining a random sample from that group?
- What is a five-number summary?
- What information about a data set does a box plot tell you? What doesn't it tell you?

5 minutes *Workshop Facilitator/Site Leader*
Homework Assignment

- Do Lessons 1 and 2.
- Review Lesson 3.
- Look at the sample student work for Lesson 3 (page 138).

BEFORE WATCHING THIS PROGRAM ...

- Make sure to have copies available of Lesson 3: Predicting M&Ms (page 122).
- Have materials for Lesson 3:
  - TI 83 calculators (1 for each pair of participants) or Table of Random Numbers
  - M&Ms (1 small bag per person — optional)
  - M&M sample data for 100 bags
  - M&M lab sheets (1 per person).
- Have materials for Lesson 1 for teachers to take home:
  - Bags of colored blocks or tiles (1 bag per pair of participants)
  - Tally charts
  - Handout for this lesson (1 per pair of participants)
  - A class chart for final predictions (to be filled in during summary)
  - Calculators (optional).
- Have materials for Lesson 2 for teachers to take home:
  - Results from Lesson 1: Mystery Bags
  - Graph paper.

SAMPLING & PROBABILITY WORKSHOP 7: DISCOVERY

Agenda for Four-Hour Workshop

20 minutes *Workshop Facilitator/Site Leader*
Introduction

Hand out the materials for Workshop 7. Discuss the following questions:
- Why is the study of Sampling & Probability important in the middle school curriculum?
- How can we introduce Sampling & Probability so that it is relevant for our students?
- How is Sampling & Probability used to solve problems?

20 minutes *Whole Group*
View Lesson 1: Mystery Bags

While watching the program, consider the following focus questions:
- What prerequisite skills do students need to be successful with this lesson?
- What is the difference between theoretical and experimental probability?
- What does it mean when the probability of an event happening is 0? 1?
- How many different ways can probabilities be stated or written?
- How many trials are necessary to be able to make an accurate prediction about an event?
- How can this lesson be used to reinforce basic skills?

30 minutes *Small Groups or with a Partner*
Read, Do and Discuss

Read Lesson 1: Mystery Bags.
Do Lesson 1.
Discuss findings, focusing on the above questions. Be sure to discuss in your group how the data helps you predict the contents of the mystery bag.

20 minutes *Whole Group*
Discuss

Discuss findings, focusing on the above questions. Make your own class chart to record each group's final predictions for the fraction of each color of tile in the bag.

20 minutes *Whole Group*
View Lesson 2: Graphing Probabilities

While watching the program, consider the following focus questions:
- What are the two variables that are being graphed?
- What type of graph (bar graph, line graph, etc.) is most appropriate for this display?
- What would be an appropriate scale for the x-axis? The y-axis?
- If you are graphing the number of trials versus the probability of an event, what would you expect to happen to the graph with more and more trials?

BEFORE WATCHING THIS PROGRAM ...

- Make sure to have copies available of all the lessons in this workshop.
- Have materials for Lesson 1:
  - Bags of colored blocks or tiles (1 bag per pair of participants)
  - Tally charts
  - Handout for this lesson (1 per pair of participants)
  - A class chart for final predictions (to be filled in during summary)
  - Calculators (optional).
- Have materials for Lesson 2:
  - Results from Lesson 1: Mystery Bags
  - Graph paper.
- Have materials for Lesson 3:
  - TI 83 calculators (1 for each pair of participants) or Table of Random Numbers
  - M&Ms (1 small bag per person — optional)
  - M&M sample data for 100 bags
  - M&M lab sheets (1 per person).

30 minutes *Small Groups or with a Partner*
Read, Do and Discuss

Read Lesson 2: Graphing Probabilities.
Do Lesson 2.
Discuss findings, focusing on the above questions. Make sure groups do line graphs. The independent variable (x-axis) is the number of trials. The dependent variable (y-axis) is the experimental probability of drawing that color.

20 minutes *Whole Group*
Discuss

Discuss findings, focusing on the above questions. Be sure to have each group share their graphs.

20 minutes *Whole Group*
View Lesson 3: Predicting M&Ms

While watching the program, consider the following focus questions:
- What is a sample and how is it used?
- What makes a sample truly representative of the larger population?
- If you have a population of 100 items (like the M&M bags), what are some methods for determining a random sample from that group?
- What is a five-number summary?
- What information about a data set does a box plot tell you? What doesn't it tell you?

30 minutes *Small Groups or with a Partner*
Read, Do and Discuss

Read Lesson 3: Predicting M&Ms.
Do Lesson 3.
Discuss findings, focusing on the above questions.

20 minutes *Whole Group*
Discuss

Discuss findings, focusing on the above questions.

10 minutes *Workshop Facilitator/Site Leader*
Homework Assignment

- Review Lesson 3.
- Review the sample student work for Lesson 3 (page 138).
- Review the Launch-Explore-Summarize Teaching Model (page 152).
- Review the Why This Topic Matters section (page 150).
- Use your journal to reflect on the focus questions from this workshop. Describe how these lessons deepened your content knowledge about Sampling & Probability. What did your learn? What else do you need to learn?

SAMPLING & PROBABILITY
WORKSHOP 7: DISCOVERY

■ Lesson 1: Mystery Bags

A. The Big Ideas

Purpose of This Lesson

The Law of Large Numbers tells us that as we gather more and more data — that is, conduct more and more trials in an experiment — the probabilities drawn from the experimental data approach the theoretical probabilities. This idea is difficult for students to grasp. This lesson and the following one provide students an opportunity to experiment to develop understanding of this concept.

Mathematical and Problem-Solving Goals

- Use probability to make predictions about an unknown population.
- Find the experimental probability and compare it to theoretical probability for the same event.
- Understand that probability models are used to predict what will happen in the long run over many trials.
- Know that any event's probability is in a range from 0 (an impossibility) to 1 (a certainty).

Connections to NCTM Standards

- Make judgments about the likelihood of uncertain events and be able to connect those judgments to percents or proportions.
- Develop conclusions about a characteristic of a population from a well-constructed sample.
- Through simulations, develop an understanding about when differences in data may indicate an actual difference in the populations from which the data were collected and when the differences may result from natural variation in samples.
- Use the language of mathematics as a precise means of mathematical expression.

B. The Lesson

Recommended Mathematical Background

- Experience converting among fractions, decimals and percents, with or without a calculator
- Some experience with graphing, ratios and proportions

Materials

- Bags of colored blocks or tiles (1 bag per pair of students) (see the Additional Notes for suggested combinations of tiles in the bags)
- Tally charts
- Handout for this lesson (1 per pair of students)
- A class chart for final predictions (to be filled in during summary)
- Calculators (optional)

TEACHING TOOLS

- As you discuss and use this lesson, make sure to take advantage of the Teacher Planning Tools (pages 151-157):
  - Planning a Math Unit: Launch-Explore-Summarize Teaching Model
  - Lesson Planner Template
  - Questions to Stimulate Student Thinking
  - Guidelines for Grouping.
- Refer to Why This Topic Matters (page 150) to help make this topic relevant to your students.
- Check out our Web site (www.learner.org/channel/workshops/missinglink) for additional tools and resources — and to join The Missing Link online discussion forum.

Time

• 60 minutes

Lesson Overview

In this investigation, students work with a partner to conduct a simple probability experiment. They draw tiles out of a bag and record the color of each draw without looking inside the bag. They then predict, after five trials, 10 trials, 15 trials and so on, up to 30 trials, the mix of colors in their bag and try to find another group that has a bag with identical contents.

Launch

Before beginning the lesson, discuss with students the meanings of probability — a way of making judgments about the likelihood of uncertain events — and sampling — a way of using a subset of a given population to gain information about the whole group. Discuss also the difference between experimental probability and theoretical probability.

Do a brief whole-class activity to quickly assess your students' understanding of probability. Take a paper bag and drop in one yellow and one red tile. Ask students if they were to pick a tile from the bag, what color it would be. Students should realize that both colors have an equal chance of showing up. How would you tell someone what the chance is of drawing a red? When they respond with 50 percent or 1/2 or one out of two, show them how to write it.

$$P(R) = 1/2 \qquad\qquad P(Y) = 1/2$$

Then have a student draw a tile and put it back in the bag at least 10 times. Let students compare the experimental results to the expected probabilities. If you get one color in a row several times, does that guarantee the other color the next time? Students will begin to see that probability is a way to explain chance but it cannot tell you exactly what will happen.

Then add two blue tiles to the bag and have students write the probabilities of each color. What is the probability of drawing a blue if the red and yellow tiles were removed from the bag? When something definitely *will* happen, the probability is 1. What is the probability of drawing a red tile? When an event *cannot* happen, the probability is 0.

Hold up a paper bag and tell the class you have put some blue, yellow, red and/or green tiles in the bags. Tell students that some of the bags have exactly the same combination of tiles. The students' job is to find out which bags are identical without looking in the bags. Tell students they should draw out one tile at a time, record the color and return it to the bag to collect data about the contents of their bag. They will make several predictions about the contents of the bag as they collect their data, and their final predictions will be recorded on a class chart. Give students the handout for this lesson.

Explore

Students work in pairs. They draw tiles from their bag one at a time and record the color for each draw. After five draws, each pair predicts what fraction of the bag is blue, yellow, red and/or green. As students continue drawing and recording tiles, predictions should be made again after 10 draws, 15 draws, 20 draws, etc.

Your primary job during this part of the lesson is to observe. Circulate among the groups, listening and watching while students work. This can give you important clues to their thinking and allow you to better focus the discussion you will lead in the lesson summary.

Your secondary job during this part of the lesson is to help students who are stuck get unstuck. Try to do this by means of leading questions; avoid "taking over" the situation.

While students work, make sure you have a class chart ready for groups to record their final predictions and to use for the class to determine which groups have identical bags.

	Group #	# blue	Fraction blue	# yellow	Fraction yellow	# red	Fraction red	# green	Fraction green
1									
2									
3									

Summarize

Challenge students to explain why they made certain predictions and to verify their predictions. When summarizing this lesson, be sure to investigate with students:

- How their predictions changed with more trials.
- How they determined their final fraction for each color.
- Which bags they think have identical contents and why.
- Did any color *not* show up? Are you 100 percent certain that color is not in the bag? Explain.
- If the color did not show up after 100 trials, can you be 100 percent certain that color is not in the bag? How many trials would it take to be 100 percent certain?

Have students share their predictions about the composition of their bags and about which bags match.

Make sure they understand the core concepts of this lesson:

- probability allows us to make predictions about an unknown population
- the probability of an event is greater than or equal to 0 and less than or equal to 1

Part of the purpose of the Summarize is to allow you to assess how well your students are progressing toward the goals of the lesson. Use the discussion to help you determine whether additional teaching and/or additional exploration by students is needed before they go on to Lesson 2.

Additional Notes

- Experimental probability is the probability determined by a simulation or experiment, such as tossing a coin many times and keeping track of the number of heads. The experimental probability is the number of heads compared to the total number of tosses. Theoretical probability is the probability obtained by analyzing a situation and listing all the possible outcomes. For example, if you toss a coin you have two possibilities: a head or a tail. Your theoretical probability of getting a head is 1 of 2.
- Certain combinations of tiles lead to richer discussion than others. Many good possibilities exist. Here is one set of combinations that works well. Make sure you make at least three bags of each combination and give them randomly to pairs of students.

Combination 1:	8 blue, 4 yellow, 12 red and 0 green
Combination 2:	4 blue, 0 yellow, 8 red and 12 green
Combination 3:	8 blue, 12 yellow, 4 red and 0 green
Combination 4:	12 blue, 8 yellow, 0 red and 4 green
Combination 5:	12 blue, 4 yellow, 8 red and 0 green

- This lesson gives opportunities for students to practice mental math computations; for instance, how would they turn the fraction 18/30 into a percentage? A decimal? We want students to use the different representations of fractions, decimals and percents interchangeably.
- Have students demonstrate that the sum of the fraction values they have predicted is equal to 1. In other words, the sum of the fractions of each color equals one whole bag of tiles.
- Students will need to save their data from Lesson 1 to use in Lesson 2: Graphing Probabilities.

MYSTERY BAGS

Your teacher has prepared bags with color tiles. Some of the bags have identical color combinations. One at a time, you and your partner will select a tile from the bag, record the color of the tile and return the tile to the bag. Repeat the procedure for 30 trials, recording the color after each trial.

A. How many tiles drawn by your group were blue? _____ Yellow? _____ Red? _____ Green? _____ Which color of tile do you think there are the greatest number of in the bag?_____ Which color tile do you think there are the least number of? _____ Explain.

B. Based on your experimental data, what fraction of tiles in the bag do you think are blue?_____ Yellow? _____ Red?_____ Green?_____ Record your findings on a class chart. Using the class information, determine which groups have the same color combination in their bags. Explain your decision.

C. Is each tile equally likely to be selected from the bag? _____ Explain. Is each color equally likely to be selected from the bag? _____ Explain.

D. What is the probability of drawing a white tile from the bag? _____ How could you add tiles to your bag so that the probability of drawing a green tile would be 1/2? (If it is already 1/2, you must add tiles and have it remain 1/2.) Be specific.

■ Lesson 2: Graphing Probabilities

A. The Big Ideas

Purpose of This Lesson

This lesson should help students make more sense out of the data from Lesson 1 by creating a visual representation (graph) of the information.

Mathematical and Problem-Solving Goals

- Understand that probability models predict what will happen in the long run over many trials.
- Use a graph to demonstrate that, as the number of trials increases, the experimental probabilities stabilize.
- Understand that, as the sample size increases, the experimental probability approaches the theoretical probability.
- Know that the probability of an event is a value equal to 0, between 0 and 1, or equal to 1.

Connections to NCTM Standards

- Choose, create and use various graphical representations of data (line plots, bar graphs, stem-and-leaf plots, histograms, scatter plots, circle graphs, and box-and-whisker plots) appropriately and effectively.
- Develop a repertoire of mathematical representations that can be used purposefully, flexibly and appropriately.
- Extend mathematical knowledge by considering the thinking and strategies of others.

B. The Lesson

Recommended Mathematical Background

- Graphing ordered pairs on a coordinate grid

Materials

- Results from Lesson 1: Mystery Bags
- Graph paper

Time

- 40–60 minutes

Lesson Overview

By graphing the probabilities of each color after each set of five draws from Lesson 1: Mystery Bags, students will create a visual model of the experimental probabilities over time. The graph will show that, as the sample size increases, the experimental probability approaches the theoretical probability. It is also important that students understand that probability is stated as a percent ranging from 0 percent to 100 percent or as decimals ranging from 0 to 1. This is reinforced as students determine the scale to use for the graph.

Launch

To graph the probabilities from Mystery Bags, have students determine the probability of each color in the bag after five trials, 10 trials, 15 trials and so on. up to 30 trials, using their data from Lesson 1. Have them plot on a

TEACHING TOOLS

- As you discuss and use this lesson, make sure to take advantage of the Teacher Planning Tools (pages 151-157):
  - Planning a Math Unit: Launch-Explore-Summarize Teaching Model
  - Lesson Planner Template
  - Questions to Stimulate Student Thinking
  - Guidelines for Grouping.
- Refer to Why This Topic Matters (page 150) to help make this topic relevant to your students.
- Check out our Web site (www.learner.org/channel/workshops/missinglink) for additional tools and resources — and to join The Missing Link online discussion forum.

coordinate grid the experimental probabilities of each color if you had stopped after five trials, 10 trials, 15 trials and so on, up to 30 trials. You might need to remind students about independent (x-axis) and dependent (y-axis) variables. Make sure they use different colors on the graph to distinguish among the results of each color tile. Ask students to be ready to explain: How did the information in the graph support your findings from Mystery Bags? How did the number of draws or trials affect your prediction? What must be true about the sum of the experimental probabilities of each color after each trial? If you conducted 30 more trials, what would you expect the graphs to look like? Explain. Reminder: If you continue drawing tiles, the probabilities should level out, and the graphs should become horizontal lines.

Then let students look in their bags. Have them find the fraction of blue, yellow, red and green tiles in the bag. Ask them to explain: How do the actual fractions compare to your experimental data and your graph?

Explore

If necessary, stop the Explore to clarify instructions that are not clear to the class. In the television studio, for example, the Learner Teachers weren't clear that their graph should include all four colors on one coordinate axis.

Students also may misinterpret the directions by graphing the number of trials versus the number of tiles drawn of each color. Unless multiple groups are graphing the wrong variables, wait until the Summarize to confront this sort of mistake. Let the group that graphed incorrectly see its mistake in the context of the class discussion.

Encourage students to see patterns in the graph. They should notice that changes in the experimental probabilities decrease with more trials.

Summarize

Have students present their findings. Make sure they understand the following core concepts of this lesson:

- Probability is an estimate of behavior over the long run.
- The accuracy of the prediction improves as the number of trials increases.
- With more trials the experimental probability becomes more stable.

Part of the purpose of the Summarize is to allow you to assess how well your students are progressing toward the goals of the lesson. Use the discussion to help you determine whether additional teaching and/or additional exploration by students is needed before they go on to Lesson 3.

Additional Notes

- Be prepared when students throw you a curve. When graphing the probability after five trials, 10 trials, 15 trials and so on, you would expect students to draw line graphs. Some students may construct bar graphs instead. Be prepared to discuss which visual representation gives the best information. A bar graph is not appropriate to show changes over time or changes that occur with more trials.
- A good extension activity would be to ask students to do 60 trials and see the extent to which the probabilities change.
- Students may try to explain the patterns they see in the graphs by determining a line of best fit. This is not appropriate. They should focus instead on how each line approaches a horizontal line with repeated trials.

GRAPHING PROBABILITIES

Use your data from Lesson 1: Mystery Bags to calculate the experimental probabilities of drawing each color after five trials, 10 trials, 15 trials and so on, up to 30 trials. Graph this data on one coordinate plane by using a different color to represent the experimental probabilities of each color tile. The number of trials is the independent variable and is graphed on the x-axis. The experimental probability is the dependent variable and is graphed on the y-axis.

Data Table for Graphing Probabilities

Trials	Trials	Trials	Trials	Trials	Trials
1.	6.	11.	16.	21.	26.
2.	7.	12.	17.	22.	27.
3.	8.	13.	18.	23.	28.
4.	9.	14.	19.	24.	29.
5.	10.	15.	20.	25.	30.

Probability after five trials:	Probability after 10 trials:	Probability after 15 trials:	Probability after 20 trials:	Probability after 25 trials:	Probability after 30 trials:
P(B) =	P(B) =	P(B) =	P(B) =	P(B) =	P(B) =
P(Y) =	P(Y) =	P(Y) =	P(Y) =	P(Y) =	P(Y) =
P(R) =	P(R) =	P(R) =	P(R) =	P(R) =	P(R) =
P(G) =	P(G) =	P(G) =	P(G) =	P(G) =	P(G) =

■ Lesson 3: Predicting M&Ms

A. The Big Ideas

Purpose of This Lesson

This lesson helps students understand the concept of random sampling. They learn how to make accurate and fair predictions — based on representative samples of data.

Mathematical and Problem-Solving Goals

- Select a random sample from a population.
- Use sampling distributions, measures of center and measures of spread to describe and compare samples.
- Use data from samples to estimate a characteristic of a population.
- Apply probability to choosing random samples of data.
- Use box plots to visualize patterns in data.

Connections to NCTM Standards

- Choose, create and use various graphical representations of data (line plots, bar graphs, stem-and-leaf plots, histograms, scatter plots, circle graphs, and box-and-whisker plots) appropriately and effectively.
- Describe and interpret the spread of a set of data using tools such as range, interquartile range, and box-and-whisker plots.
- Interpret graphical representations of data, including description and discussion of the graph's shape and features of the graph, including symmetry, skew and outliers.
- Find, describe and interpret mean, median and mode as measure of the center of a data set; know which measure is best to use in particular situations; and understand how each does and does not represent the data.
- Develop conclusions about a characteristic in the population from a well-constructed sample.
- Extend mathematical thinking by considering the thinking and strategies of others.
- Develop a repertoire of mathematical representations that can be used purposefully, flexibly and appropriately.

B. The Lesson

Needed Mathematical Background

- Understanding the median of a set of data

Materials

- TI 83 calculators (1 for each pair of students) or Table of Random Numbers
- M&Ms (1 small bag per person — optional)
- M&M sample data for 100 bags
- M&M lab sheets (1 per person)

Time

- 60–80 minutes

TEACHING TOOLS

- As you discuss and use this lesson, make sure to take advantage of the Teacher Planning Tools (pages 151-157):
  - Planning a Math Unit: Launch-Explore-Summarize Teaching Model
  - Lesson Planner Template
  - Questions to Stimulate Student Thinking
  - Guidelines for Grouping.
- Refer to Why This Topic Matters (page 150) to help make this topic relevant to your students.
- Check out our Web site (www.learner.org/channel/workshops/missinglink) for additional tools and resources — and to join The Missing Link online discussion forum.

BORROW FREE CALCULATORS

- Texas Instruments has a calculator loan program; see its Web site at www.ti.com/calc/docs/loan.html. The loan program is for teachers and teacher workshops, but if calculators are available, TI will loan them for classroom use, too. The Web site has lots of teacher resources. Another site of interest is Teachers Teaching with Technology (T3) at www.t3ww.org/t3. This site also has great technology ties for teachers, including training opportunities.

Lesson Overview

This investigation introduces students to the concept of random sampling. For a sample to be random, all members of the population must have an equally likely chance of being selected. In this lesson, students use calculators to generate random numbers from 1 to 100 to select samples of 25 M&M bags from a previously collected database of 100 bags. They then analyze samples using box plots to help them predict what percentage of each color M&M the candy company produces.

Launch

First, you'll need to teach students how to construct box plots (see mini-lesson, page 124) and how to use graphing calculators, if they're available. (See sidebar on page 122 and discussion on page 125.)

Once you've done the box-plot mini-lesson and shown students how to conduct a random sample, introduce the problem (see handout, Predicting M&Ms, page 128).

What Percent of Each Color of M&M Does the Candy Company Manufacture?

Make sure students understand all of the components of the task. Part of this investigation is to be done individually by students and other parts completed by the group.

Individually:

- Select a random sample of 25 bags from the M&M database.
- Complete the five-number summary for each color.
- Make a box plot for each color on separate record sheets for the group.
- Summarize the group findings by explaining how the predictions were made and how they were determined based on the box-plot representations.

As a Group:

- Discuss the results of each random sample.
- Analyze the box plots for each color.
- Determine how to use the results to predict the actual percent of each color M&M the candy company produces.
- Make predictions.

Explore

Make sure each student has a copy of the M&M database and all the handouts that go with the problem. Students investigate this problem individually and then statistical data is shared in a group to come up with a final prediction for each color of M&M.

Even if students are working independently, let them check with a small group if they're having trouble. This can be a very powerful learning experience. As one Learner Teacher said after the workshop, "During this lesson, I went from 'I think I can't' to 'I know I can.' And if I can't, I know I have two people at my table who can help."

Summarize

Have groups report their findings, making sure the students understand the core concepts of this lesson:

- Samples can be used to make predictions about a large population.
- Random sampling provides more representative data about a population than nonrandom sampling.
- Box plots organize data and enable us to estimate a characteristic of a population.

Help students see the visible patterns in their data and graphs. Where are the points tighter and more spread out? What does the size of the box tell me? Are the samples similar? How confident am I of my predictions, based on the box plots our group has constructed?

If you have real bags of M&Ms, make sure to let students count them at the end, compare the actual number of each color with their predictions — and then eat them!

Part of the purpose of the Summarize is to allow you to assess how well your students are progressing toward the goals of the lesson. Use the discussion to help you determine whether additional teaching and/or additional exploration by students is needed before they go on to future topics using The Missing Link model.

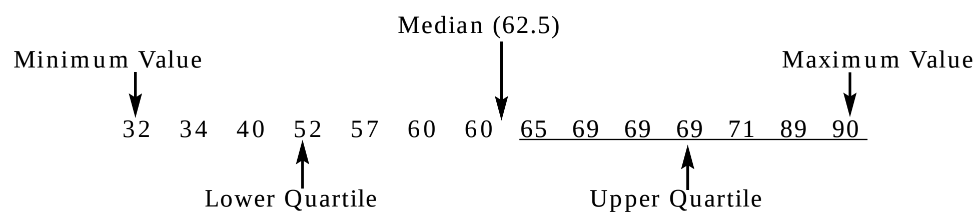
Additional Notes

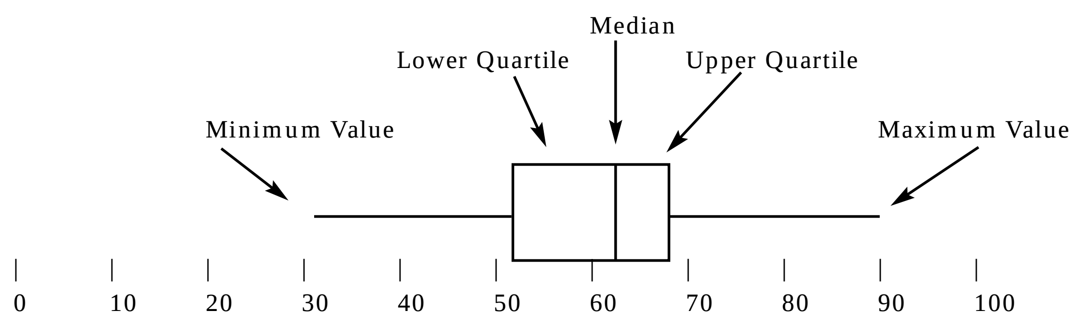
- This lesson is a real equalizer — a good place for students of differing abilities to work together.
- The mini-lesson or refresher lesson on box plots is a good example of when to do some extended direct instruction to introduce or reinforce a particular concept. Your students will be more likely to understand box plots because they're learning them to immediately solve a problem.
- Students really might struggle with this lesson. It's difficult math. Be careful not to say, "Come on guys, you can do this. It's easy!" Think about the child who doesn't get it — he or she will really begin to doubt his or her own capabilities. Students must learn that struggling with a concept or a problem is acceptable. The person who struggles, perseveres and then understands the problem is the winner.
- Take advantage of middle school students' keen interest in themselves — if they need more practice with box plots, have them plot samples of their classmates' average heights, the amount of TV they watch or the free-throw percentages of members of the basketball team.

Mini-Lesson: Constructing Box Plots

A box plot is constructed from the five-number summary of the data. The five-number summary includes the minimum value, the lower quartile, the median, the upper quartile and the maximum value. The minimum and maximum values are simply the smallest and largest numbers in the data set. The median is the number that separates an ordered set of data in half. The lower quartile is the median of the data values below the median. The upper quartile is the median of the data values above the median.

The example below shows how the five-number summary corresponds to the features of the box plot.





This display does not give actual data points, but it does tell us about the distribution of the values. The “box” contains 50 percent of the data — 25 percent of it below the median line and 25 percent above the median line. If the box is small and tightly compacted, it tells the reader that half of the data falls in a very small range. If, on the other hand, the box is spread out, the data must cover a wide range of values. Each “whisker” contains another 25 percent of the data.

With some data sets, you may encounter outliers, data values much greater than or less than most of the other values in the set. To determine if a value is an outlier, first calculate the interquartile range (IQR), the difference between the upper and lower quartiles. Then outliers are determined as follows:

- Values > 1.5(IQR) + upper quartile
- Values < lower quartile – 1.5(IQR)

Outliers are generally a few values only; the actual box and whiskers show the spread of most of the data. Looking at the data without the outliers makes it easier to analyze and make generalizations about the data without being misled or distracted.

Extension — The Effect of Sample Size

An extension of this investigation explores how the size of a sample affects the accuracy of statistical estimates. Focusing only on the red M&Ms, students select three random samples of five bags and find the median number of reds for each sample. All the median values then are collected on a class line plot. The same procedure is completed for a sample size of 10 bags and 25 bags (data from the first part is used for the 25-bag sample). Students then can compare the line plots to determine if sample size has any effect on the median values. Generally, the larger sample size will show a clear clustering of median values, while the smaller sample size will create a much greater spread on the line plot. This activity helps to reinforce the concept that the greater the sample size, the more representative the sample is of the entire population.

Random Numbers on the TI 83

One key ingredient in simulations is the idea of randomness — tossing a coin, rolling a die, drawing a marble out of a bag. Although simulations can involve such things as choosing a number from a hat, it is also possible to do the simulations on a calculator. To generate random numbers on the TI 83, the random numbers must be “seeded” with a starting value; make sure all students use a different starting value. If they are not seeded but left to the default, then all students will generate the *same* “random” sequence. In doing simulations in the classroom, it is essential that all students generate their own unique set of random digits. To seed a random generator, follow the keystrokes:

[(input any number)] [STO] [MATH] [→] [→] [→] PRB [ENTER] [ENTER]

Once the calculator is seeded with a starting value, random numbers can be generated by the following sequence of keystrokes:

[MATH] [→] [→] [→] PRB [ENTER] [ENTER]

To generate random integers (the equivalent of rolling an n-sided die), simply follow the keystrokes below:

[MATH] [→] [→] [→] PRB [5: randInt] [1] [,] [6] [,] [1]

DATA FROM 100 BAGS OF PLAIN M&MS

Bag	Green	Yellow	Orange	Blue	Brown	Red	Total
1	3	10	9	5	10	18	55
2	5	12	4	6	19	11	57
3	7	10	9	4	16	12	58
4	4	14	2	1	14	19	56
5	12	7	8	7	14	13	61
6	10	9	6	5	15	8	55
7	11	11	6	6	12	12	58
8	8	15	5	3	16	10	57
9	2	11	4	4	24	12	57
10	5	7	4	1	26	13	56
11	6	13	4	4	15	18	60
12	5	8	4	2	23	16	58
13	9	13	4	4	14	11	55
14	9	10	5	5	14	14	57
15	5	19	5	2	13	14	58
16	3	15	5	2	19	11	55
17	3	10	4	3	23	14	57
18	6	7	5	5	15	22	60
19	5	7	3	4	21	14	54
20	8	7	8	2	20	16	61
21	10	11	7	7	8	14	57
22	7	10	3	5	20	12	57
23	3	8	6	3	25	10	55
24	6	11	9	3	10	17	56
25	10	12	1	2	15	17	57
26	4	12	4	7	14	16	57
27	6	9	6	7	15	13	56
28	5	11	6	7	17	7	53
29	1	10	6	5	22	14	58
30	10	4	8	0	26	9	57
31	4	14	6	4	18	12	58
32	6	18	2	4	19	14	58
33	6	7	8	4	20	11	56
34	12	11	6	4	11	11	55
35	5	10	6	2	12	16	51
36	8	9	4	4	16	17	58
37	2	12	2	6	11	21	54
38	5	7	3	4	21	19	59
39	8	7	8	2	20	16	61
40	10	11	7	7	8	14	57
41	7	10	3	5	20	12	57
42	3	8	6	3	23	10	50
43	6	11	9	3	10	17	56
44	10	12	1	2	15	17	57
45	5	13	2	4	22	11	57
46	6	10	9	5	14	13	57
47	6	16	7	3	16	9	57
48	6	10	4	5	23	10	58
49	10	7	2	6	19	9	53
50	4	12	8	6	10	15	55
51	9	9	6	6	17	10	57

Bag	Green	Yellow	Orange	Blue	Brown	Red	Total
52	4	13	4	6	17	13	57
53	6	12	3	8	13	12	54
54	11	8	8	12	9	8	56
55	1	16	7	3	22	10	59
56	6	11	6	4	19	11	57
57	7	7	7	3	10	21	55
58	7	2	8	10	15	13	55
59	6	10	6	7	12	15	56
60	6	16	7	3	16	9	57
61	6	10	4	5	23	10	58
62	10	7	2	6	19	9	53
63	4	12	8	6	10	15	55
64	9	12	8	6	8	15	58
65	10	6	5	4	12	16	53
66	4	11	3	2	21	15	56
67	6	15	4	8	10	10	53
68	6	8	7	1	19	14	55
69	6	8	8	6	10	16	54
70	9	11	7	4	15	10	56
71	6	9	8	2	19	14	58
72	3	10	9	5	10	18	55
73	5	12	4	6	19	11	57
74	7	10	9	4	16	12	58
75	4	14	2	1	16	19	56
76	1	8	10	1	22	14	56
77	5	15	4	9	11	11	57
78	3	11	6	3	24	10	57
79	10	9	4	1	23	10	57
80	5	10	7	1	21	13	57
81	6	14	7	7	14	5	53
82	9	11	2	6	13	16	57
83	7	7	9	0	13	20	56
84	8	10	4	5	13	10	50
85	4	11	2	1	24	15	57
86	4	12	6	3	21	12	58
87	5	8	7	4	20	13	57
88	7	11	7	7	13	10	55
89	9	11	4	2	12	18	56
90	4	15	8	4	16	10	57
91	7	11	6	4	18	11	58
92	5	8	8	3	20	12	56
93	7	3	2	6	26	11	55
94	9	6	3	1	28	12	59
95	12	11	9	2	18	5	58
96	9	11	3	3	17	12	55
97	5	12	6	5	17	13	58
98	4	11	9	3	21	10	58
99	11	12	5	3	17	9	57
100	6	16	6	6	16	4	54

PREDICTING M&MS

What is the Color Distribution of Plain M&Ms?

Through your investigation you must determine what percent of each color of plain M&Ms the candy company manufactures.

- Select a random sample of 25 bags (use number tiles in a bag, a random number table, a graphing calculator or some other method to conduct your random sample). For each sample bag identified, record the number of each color of M&M found in that bag on your record sheet. Note: Make sure that each M&M bag in your sample is a different number bag. That means that if you randomly select bag number four and bag number four is chosen a second time, ignore it the second time and select again.

- Calculate the five-number summary for the M&M color data for your sample.

- Make box plots of the M&M data on the group's labsheets. There should be one labsheet per color of M&M for each group.

- Once each member of the group has constructed his or her box plot for each color on the group's chart, predict the percentage of each color of M&M that is produced by the candy company.

- Explain in detail how you determined your answer. Use the mathematical representations as support and refer to them in your analysis. In other words, what do the box plots tell you about the number of each color of M&M that is being produced?

Source: *Connected Mathematics Project — Samples and Populations*

M&M SAMPLE: ANALYSIS OF COLORS

Green																								
Yellow																								
Orange																								
Blue																								
Brown																								
Red																								
Total																								

Use the data above to determine the five-number summary for each M&M color

Color	Minimum Value	Lower Quartile	Median	Upper Quartile	Maximum Value	Outliers
Green						
Yellow						
Orange						
Blue						
Brown						
Red						

M&M BOX-AND-WHISKER PLOTS

M&M Color ____________________

Sample 1

0 2 4 6 8 10 12 14 16 18 20 22 24 26 28

Sample 2

0 2 4 6 8 10 12 14 16 18 20 22 24 26 28

Sample 3

0 2 4 6 8 10 12 14 16 18 20 22 24 26 28

Sample 4

0 2 4 6 8 10 12 14 16 18 20 22 24 26 28

Prediction:

Number of Color Per Bag____________________ Percent of Color Produced____________

M&M SAMPLE: ANALYSIS OF COLOR

What is the Color Distribution of Plain M&Ms?

- What percent of each color of plain M&Ms does the candy company manufacture?

yellow ________

orange ________

blue ________

brown ________

red ________

green ________

- Explain in detail how you determined your answer. Use the mathematical representation as support. In other words, what do the box plots tell you about the number of your color M&M that is being produced?