Chunk It!

Wendy Weber
Central College
weberw@central.edu

Published: October 2016

Overview of Lesson

In this lesson, students will investigate whether grouping letters into recognizable chunks, "chunking", aids memorization via a two-sample randomization test. Students will randomly be given a sequence of letters to memorize and recall. Upon collecting the data, they will analyze it, and interpret it to determine if the data suggests "chunking" aids memorization. Students will then be introduced to a two-sample randomization test where they will simulate randomization of the data to determine the likelihood of obtaining results that are as extreme (farther from zero) as that obtained in the class experiment.

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, and interpret results in the context of the original question.
This is a **GAISE C** 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.

Learning Objectives Alignment with Common Core and NCTM PSSM

Learning Objectives	**Common Core State Standards**	**NCTM Principles and Standards for School Mathematics**
Students will learn the concepts of a two-sample randomization test, conduct tactile and computer simulations, and interpret this probability in the context of the problem.	**S-IC.5.** Use data from a randomized experiment to compare two treatments; use simulations to decide if differences between parameters are significant.	**Grades 9-12** Develop and evaluate inferences and predictions that are based on data: use simulations to explore the variability of sample statistics from a known population and to construct sampling distributions.

Prerequisites

Students will be familiar with creating & comparing dotplots, calculating means, and analyzing and interpreting graphical and numerical displays of data. Students should be familiar with the three letter groupings (words and acronyms) of the sequence of letters. English language learners may find this task challenging (see the Differentiation section for ideas on how to address this problem).

Time Required
Three 50-minute periods.

Materials and Preparation Required
One stop-watch or timer with seconds. Each student will need a calculator, a piece of paper with one of two sequences of letters, a writing utensil, and student activity pages. Each pair of students will need three sticky notes (the small square size works best) and a brown paper lunch bag filled with blank slips of 1.5”x1.5” red and blue square paper; card stock paper works best.

Each pair of students will also need an Adobe Flash-enabled computer with Internet to access the Rossman/Chance Randomization Test applet at
http://www.rossmanchance.com/applets/randomization20/Randomization.html

Chunk It!
Teacher's Lesson Plan

This lesson has two parts. The first part focuses on class level data, and the second part takes the results from Part I and extends it to investigate whether this result could have happened by chance using a two-sample randomization test. The idea for this lesson plan was derived from Chance (2105).

Part I. A Class-Level Experiment

Describe the Context and Formulate a Question

Ask students to consider what kinds of information they might memorize and what makes the information easy or more challenging to memorize. Examples where grouping, or chunking, information to make memorizing and recalling easier include phone numbers, social security numbers, computer passwords, birthdates, etc.

Consider the following question:
When asked to memorize a string of letters, how many letters can someone memorize and recall, and does it matter how we group the letters?

To investigate this question, ask the class to think about how we might try to measure memorization and recall skills. Ask students what kinds of questions we could ask to test memorization and recall. Answers will vary, but guide students to the general set-up of giving students a sequence of letters they will be asked to memorize (within a time limit such as 20 seconds) and recall.

Collect Data

Next, inform the class that we are going to conduct an experiment in which everyone will receive a sequence of letters to memorize. After 20 seconds, they will be asked to recall the information in the exact order it was given by writing it down. At this point, do not allude to the fact that there will be two different lists to memorize. This, and the main point of the activity, which is to determine whether chunking (grouping letters into recognizable chunks) aids memorization, will come out more explicitly as the activity progresses.

In preparation for the activity, cut the two sequences (in the Appendix) into slips of paper. Once the slips of paper are cut, alternate a slip of paper with the CAT list with one that has the CATF list before randomly passing them out to students. It is recommended that the data be printed onto colored paper that cannot be easily seen through.

Before collecting the data, outline the procedure for the students. The instructor will randomly pass out the two types of data collection strips (in the Appendix) by placing each face down on a student's desk. Then, when everyone is ready, students will turn over their slip of paper and the instructor will begin the stop-watch. Students will have 20 seconds to memorize the sequence of letters in the order given. When 20 seconds are over, students will turn the slip of paper over and write down the sequence of letters *in the order* they were given.

After students have completed the task, have them count the number of letters they got correct beginning on the left of the sequence. *As soon as a letter is not correct*, students stop counting. Have them write their count next to the sequence they wrote.

Next, pass out the activity sheets and ask students to talk with one another about the task and answer question #1. Students should realize that everyone received the same sequence of letters, but they were grouped differently.

After a few minutes, show the students the two sequences of letters they were asked to memorize and ask them to answer question #2 independently before sharing responses with the class.

CAT – FBI – USA – NFL – LOL – ABC – CBS – ACT – GPA – OMG
CATF – BIU – SAN – FLLO – LAB – CCB – SA – CTG – PAO – MG

Project a class table (question #3) onto an overhead screen (using a document camera or a computer projection system). Then, have students add their letter counts and the first grouping of letters in their sequence to the class table by either hand writing (if using a document camera) or typing their response (if using a computer projection system).

Analyze Data

Once students have copied the class data, have them work in pairs to create stacked dotplots, and describe the distributions using center and spread. They should calculate means and the difference between the means. Before moving on, bring the class together to talk about their responses to #4-7.

Interpret Data

Students will likely conclude that the data suggests people who were given the CAT grouping appear to recall more letters. The mean difference is likely to be greater than zero, and the two dotplots should reveal that those with the CAT grouping have more data points further to the right than those with the CATF grouping. If the mean difference is less than zero or zero, then students may conclude that the grouping method didn't aid memorization.

PART 2: Could this Have Happened Just by Chance?

Formulating a new question

Following question #7, hold a class discussion about how our data may *suggest* that forming recognizable chunks, or "chunking" aids memorization. Is it possible that we could have obtained the results we did by randomly assigning the number of letters memorized to the CAT group or the CATF group? This is our new investigative question.

We want to able to examine whether it was just by chance that people who had the CAT grouping memorized more letters. Ask how we can test that we didn't just get lucky. If chunking really has an impact on number of letters memorized, and our results did not just happen by chance, then we should obtain strong evidence that allows us to state that it's *likely* chunking aids memorization.

The process we will follow is a two-sample randomization test. We have two samples, the CAT group and the CATF group. Randomization allows us to answer the question, could the results that we obtained have been due to chance alone? We start with the hypothesis that they were due to chance alone and look for evidence to the contrary. That is, we look for evidence that says more than just chance caused the results we obtained. So, we ask the most general question, could the results we obtained be due to chance? That is, does it matter which letter grouping a person had? If it really didn't matter which group a letter count came from, then both groups would have the same mean and the mean difference would be zero. If chunking does aid in memorization, then the statistic $\bar{x}_{CAT} - \bar{x}_{CATF}$ should be positive.

Collect Data

To answer our new question, we need to collect data by considering the mean differences of randomly assigned groups under the hypothesis that there is no difference in means. So, we are going to mix up all the counts of letters memorized that we obtained in our original experiment and randomly assign each to either the group CAT or CATF. Then we are going to calculate a new mean difference. We will repeat this process many times, creating a class dot plot of the by-chance data. We want to know how likely it is that we would get a results like we did in #6 if everything was by chance alone. So, we see how often this randomization produces a mean difference at least as extreme (farther from zero) as the one we observed in #6. If the proportion of mean differences is very small, then it is unlikely that our result came from the by-chance distribution. If it didn't come from the by-chance distribution (our dot plot), then there likely was something more going on. This "something more" is the grouping of the letters. That is, the hypothesis that chunking aids memorization is a plausible explanation for the difference in means.

If we obtain many results as extreme (farther from zero) as we did in our sample, then we will know that we could have obtained results like we did just by chance (randomization) alone, and we have evidence that chunking doesn't really aid memorization. If we do not obtain many results as extreme (farther from zero) as we did in #6, then it's unlikely that the random process yielded the mean difference that we obtained; that is we have evidence that chunking aids memorization.

To understand the process of randomization, each pair of students is going to perform three simulations and pool their data with that of their classmates. If class sizes are large, it may be sufficient for students to only do two simulations by hand. The goal for the number of repetitions is to have enough data points to be able to detect overall trends in the distribution; at least 30 should give good results in the class dot plot that will be created.

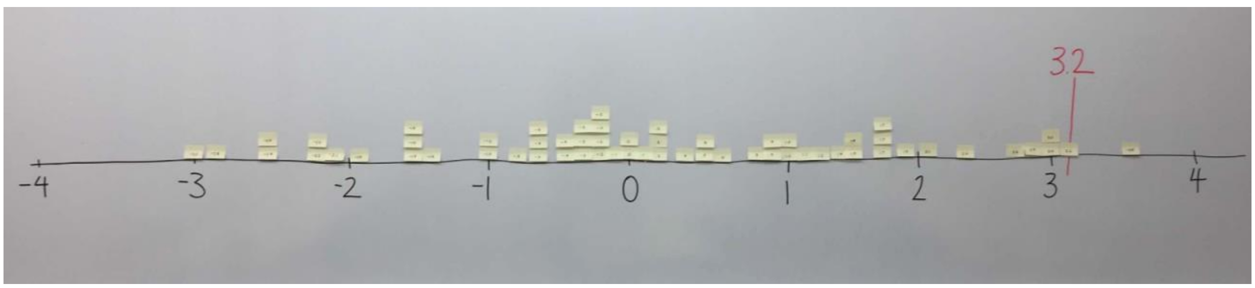
Once students have completed #9, hold a class discussion to make sure they understand what they just did. Ask students to explain the process as well as the purpose of randomization. The purpose is to see whether how many letters someone memorized correctly was the effect of the grouping they were given. The question we are trying to answer is whether the groupings could have been obtained by chance. So, we consider what happens if the groupings were randomly assigned (which is what we did by mixing up the letter counts and re-assigning them to a group).

Then we computed the mean number of letters memorized for each new randomized group and look at the difference of means between the two groups. Next, we will plot the class data of mean difference to create a by-chance distribution of mean differences. This distribution is what we would expect to get if there wasn't a difference in the grouping methods.

Most pairs of students will likely have a mean difference that is not more extreme than the observed mean difference in #6. If no one has a value more extreme than that in #6, talk about whether it *could* happen.

Analyze Data

While students are recording their mean differences on their sticky notes, create space on the board for them to add their data. The axis should be labeled at well-spaced unit intervals (big enough for sticky notes to be placed at decimal intervals) from about -4 to 4. The dot plot should be roughly symmetrical with a mean of zero (because we are assuming there is no difference between the groups). After everyone has added their sticky notes to the board, draw and label a vertical line with the observed difference in means (from #6). See below for an example using the class data provided in the solutions.



Ask the class what it means for a rearranged mean difference to be more extreme (farther from zero) than the observed difference. Since we are looking for values that are more extreme than what we observed, we want to know how many data points lie to the right (assuming the class mean was positive) of the line you drew.

Interpret Results

The next two questions guide students to compute the probability, p, of obtaining results as extreme as the ones we did (for those familiar with terminology, this is the simulated p-value) by counting the number of values at least as extreme as our observed difference and dividing by the total number of "dots"/sticky notes.

In #11 and #12, students interpret their probability in the context of the problem. If p is sufficiently small, students should conclude that there is significant evidence that it is not likely that our result from #6 didn't come from the by-chance distribution. That is, we do not have evidence to support this hypothesis, so we conclude that it is likely that chucking does aid memorization. If p is large, then it is likely that our observed value from #6 came from the by-chance distribution, that is, we have evidence that our hypothesis is plausible and that chunking likely does not aid memorization. In March 2016, the American Statistical Association (A.S.A.) released a statement saying that we should not teach students a hard and fast rule about how small p should be to determine what constitutes significance, but rather teach them to interpret p in the context of the problem. Therefore, hold a discussion about how the size of p determines whether we say that chunking is a plausible explanation. In general, the smaller p is, the stronger the evidence *against* our explanation that chunking does *not* aid memorization. Remember that we are asking whether we could have obtained results like we did (or more extreme) by chance alone. If the value of p is small, then it is not likely that we would have obtained results like we did in #6 by chance alone and there is something more going on. The only difference in the two methods was the chunking, thus, it is plausible that chunking does aid memorization.

Bring the class together to talk about the difference between finding the actual probability and estimating the probability. To find the actual probability, we would need to look at all possible rearrangements into groups the same size as in our experiment, compute the difference in means, and then count how many of these differences are at least as extreme as our observed difference. If our original CAT and CATF groups were of size 10 and 9, there would be $\binom{19}{9} = \frac{19!}{9!10!} =$ 92,378 such ways, and we would have to calculate all 92,378 possible differences in means. Whew! We found a sample of those differences and counted the observations that are at least as extreme as our observation in #6. If we wanted more confidence in our results, we would need to increase the number of randomization simulations (Tintle, 2014).

Collecting Analyzing and Interpreting with a Technology-based Simulation

To increase the number of randomization simulations, we would either need to have each pair of students conduct more tactile simulations, or we would need to use technology to do so. The student handout contains more detailed directions and questions 13-16 for students. Fortunately, there are applets that do what we did by hand quite quickly. The Rossman/Chance Randomization Test applet at http://www.rossmanchance.com/applets/randomization20/Randomization.html is a nice applet that exactly imitates the tactile simulation if you do one repetition. Students will first have to enter the data into the applet for conducting a simulation. They will first perform one repetition, then an additional 20, for a total of 21 repetitions. Finally, they will rest the app with their data and perform 1000 repetitions. Using the 1000 repetitions, they will use the applet to calculate the number of randomizations that are at least as extreme as the value they got in #6 and calculate the corresponding percentage of 1000. As they did in the tactile simulation analysis, students will determine whether it is likely that the results we obtained in #3 and #6 were due to chance (large value of p) or due to chunking (small value of p).

To close the activity, return to the examples generated by students at the beginning of class. How many of the examples given used chunking to aid memorization? Are there other techniques that aid memorization? Using visualization (mentally visualizing the letters) is another powerful technique to aid memorization.

Suggested Assessment

In the assessment that follows, teachers should note that we are moving from using $\bar{x}$, which denotes the sample mean, to using μ, which represents the (often unknown) population mean.

The questions that we ask are asking students to use the given data and make inferences about the populations from which the data are taken.

In a survey of 11th graders in 2014 (http://www.amstat.org/censusatschool), students were asked to rate how important recycling is to them. Students were presented with a sliding scale from 0 (not important) to 1000 (very important). The results of random samples of students in Iowa and from Colorado are given in the tables below.

IOWA Importance of Recycling	200	225	400	455	487	500	500	500	540	583
	599	600	600	643	749	845	904	945	950	1000
	1000	1000								

COLORADO Importance Recycling	0	70	200	400	500	543	583	588	600	600
	600	700	700	700	700	749	750	754	797	800
	800	810	905	950	1000					

Do 11th grade Iowa students rate the importance of recycling higher than 11th grade students from COLORADO?

1. Construct stacked dotplots of the two data sets. Draw a vertical line on each showing the mean importance rating.

Let μ_{IA} be the population importance level of 11th grade Iowa students, and let μ_{CO}be the population importance level of 11th grade Colorado students.

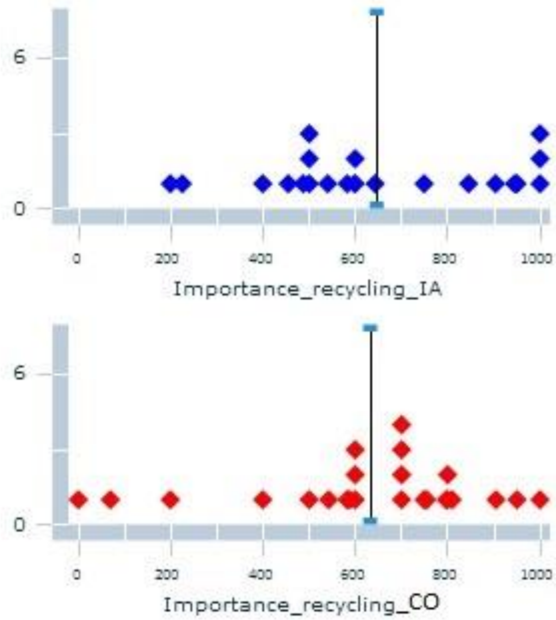
2. If we believe that Iowa students rank recycling higher than Colorado students, then would you expect the difference $\mu_{IA} - \mu_{CO}$ to be positive, negative, or close to zero? Explain your answer.

3. Jules performed 40 re-randomizations and computed the difference $\bar{x}_{IA\_rearranged} - \bar{x}_{CO\_rearranged}$ for each randomization. Her collection of re-randomized mean differences is given below. Based on this data, is there evidence that 11th grade Iowa students rank recycling higher than 11th grade Colorado students? Explain your answer.

4.205	-76.976	-16.816	51.12	-6.647	-1.776	-76.378	-24.935
-18.269	-69.371	91.284	143.924	-42.367	-19.038	106.495	166.142
-107.398	-27.156	31.124	80.858	112.391	-15.278	45.907	16.767
24.116	90.002	49.582	-124.916	97.864	136.062	93.249	122.218
19.587	106.495	-117.567	-30.318	-144.315	40.267	-49.631	-132.607

Assessment Solutions

1.



$\bar{x}_{IA}$ = 646.59
$\bar{x}_{CO}$ = 631.96
$\mu_{IA} - \mu_{CO}$ = 646.59 – 631.96 = 14.63

2. $\mu_{IA} - \mu_{CO} > 0$. Iowa students rate recycling as more important than Colorado students so μ_{IA} will be larger than μ_{CO}.
3. Based on this data (count the number of re-randomized mean differences greater than $\mu_{IA} - \mu_{CO}$ = 14.63 and divide by the total number of data points), the estimated p-value is $\frac{20}{40}$ = 0.5. Thus, there is no evidence that 11th grade Iowa students rate the importance of recycling higher than 11th grade Colorado students.

4.205	-76.976	-16.816	51.12	-6.647	-1.776	-76.378	-24.935
-18.269	-69.371	91.284	143.924	-42.367	-19.038	106.495	166.142
-107.398	-27.156	31.124	80.858	112.391	-15.278	45.907	16.767
24.116	90.002	49.582	-124.916	97.864	136.062	93.249	122.218
19.587	106.495	-117.567	-30.318	-144.315	40.267	-49.631	-132.607

Possible Differentiation

To assist English Language Learners, the instructor might revise the activity to include more three letter words or acronyms they would be familiar with. For instance, one such sequence of letters might be CAT-THE-DOG-AND-SHE-ABC-ALL-DAY-HAS-CAN. Furthermore, http://www.rollingr.net/wordpress/2007/02/02/common-letter-sequence/ has some useful information about which letters in the English language occur most frequently, which should aid ELL students when using the list above. Students can extend the ideas in this activity using the most common tri-graphs (three letter groupings in the English language) or most common three-letter words in a letter sequence to determine whether chunking aids memorization when the words are more common.

References

Census at School - United States. (n.d.). Retrieved December 16, 2015, from http://www.amstat.org/censusatschool/index.cfm

Chance, B., & Rossman, A. (2015). *Investigating statistical concepts, applications and methods, second edition.* Belmont, California: Duxbury.

Randomization. (n.d.). Retrieved December 16, 2015, from http://www.rossmanchance.com/applets/randomization20/Randomization.html

Tintle, N., & Chance, B. (2014). *Introduction to statistical investigations* (Preliminary ed.). Hoboken NJ: Wiley.

Further Reading About the Topic

Tintle, N., Topliff, K., VanderStoep, J., Holmes, V., Swanson, T. (2012). Retention of Statistical Concepts in a Preliminary Randomization-Based Introductory Statistics Curriculum. *Statistics Education Research Journal*, 21-40. http://iase-web.org/documents/SERJ/SERJ11(1)_Tintle.pdf

Chunk It!
Student Handout

PART 1: Analyzing our Class Data

So, how good is your memory? Some students memorized more letters than others. Are they better at memorizing or is something else going on?

1. Compare the letters you were asked to memorize with those of others in class. What is different about the sequence of letters?

2. Which group would you expect to memorize more letters? Explain your answer.

3. Count the number of letters you correctly recalled. Starting on the left, count the letters that are exactly the same as the given sequence. As soon as a letter is incorrect, stop counting. For example, if you had the grouping CATF-BIU-SAN-FLLO … and you wrote down CATF-BUI-SAN-… , you would record a count of 5. Letters must be in order!

   Record your count in the class data sheet. When all students have entered their data, copy it to the table below.

# Letters Correctly Recalled	CAT or CATF Group	# Letters Correctly Recalled	CAT or CATF Group	# Letters Correctly Recalled	CAT or CATF Group	# Letters Correctly Recalled	CAT or CATF Group

4. For each group (CAT or CATF) of data, create a dotplot. Your plots should be well-labeled, have the same scale, and be such that one is stacked above the other for easy comparison.

5. Describe the shape, center ($\bar{x}_{CAT}$ or $\bar{x}_{CATF}$) and variability of each dotplot.

6. We can also compare the two centers by computing the observed difference in the means:

$$\bar{x}_{CAT} - \bar{x}_{CATF} =$$

7. Consider your answers from the previous two questions. Does the data *suggest* that one grouping of letters is easier to memorize than the other? Explain your answer.

<u>PART 2: Could this Have Happened by Chance?</u>

Does chunking (grouping letters to form recognizable chunks of letters) impact the number of letters people can correctly recall, or did we just get lucky with our sample?

To answer this question, we will perform a *two-sample randomization test*. The purpose of this test is to determine whether the results that we obtained in #3 and #6 are really due to chunking or if they could have happened just by chance. If there really were no difference between groups, then we would expect the counts for the two groups to be the same and the mean difference to be 0. So, we start with this assumption. We will investigate the following question: Could the results we obtained have occurred just by chance?

To see whether the number of letters memorized was due to chance, we are going to randomly assign the number of letters memorized to either the CAT or CATF group. To do this, we mix up all the counts of letters memorized and randomly deal each to either the group CAT or CATF.

Then we are going calculate a new mean difference. We will repeat this process many times, and then see how often this randomization produces a mean difference at least as extreme (farther from zero) as the one we observed in #6.

If we obtain many results as extreme (farther from zero) as we did in #6, then we will know that we could have obtained results like we did just by randomization alone, and chunking doesn't really aid memorization. That is, the results that we saw in #6 could have occurred just by chance alone. If we do not obtain many results as extreme (farther from zero) as we did in #6, then we know that it wasn't just by the random process alone that we were able to obtain a mean difference like we did; that is chunking really aids memorization.

To understand the process of randomization, each pair of students is going to perform three simulations by hand. Then we are going to pool our results and see how often we obtain results like we did in #6.

Before carrying out the simulation, read all directions!

To keep track of which group our data belonged to originally, we are going to use colored paper. All the data from the CAT grouping should be written on red slips of paper. All the data from the CATF grouping should be written on blue paper.

Simulation:

a. On the correct color (red or blue) of slip of paper, write the number of letters you memorized.
b. Repeat this process for every data point collected by the class (see #3).
c. Next, put all the slips of paper into a brown bag and shake it up.
d. Without looking, draw a slip of paper and assign it to pile A.
e. Draw a second slip and assign it to pile B.
f. Continue to alternate which pile receives a slip of paper until pile A has the same number of data points that our CAT group did. That is, Pile A is our random assignment of people to the CAT chunking group (A = re-randomized CAT) and Pile B is our new CATF (B = re-randomized CATF).
g. Calculate the following statistics:

$\bar{x}_A = \bar{x}_{\text{re-randomized } CAT} =$ $\qquad$ $\bar{x}_B = \bar{x}_{\text{re-randomized } CATF} =$

$$\bar{x}_{\text{re-randomized } CAT} - \bar{x}_{\text{re-randomized } CATF} =$$

8. Is this difference from your first randomization more extreme (farther from zero) than the observed mean difference given by the data (#6)?

9. Complete the randomization two more times and record your mean differences in the table below.

	Trial 1	**Trial 2**	**Trial 3**
Mean difference: $\bar{x}_{\text{re-randomized } CAT} - \bar{x}_{\text{re-randomized } CATF}$			

You have just completed three simulations. If we repeated this simulation many times, how often would we obtain values as extreme (farther from zero) as that in #6?

Instead of each group doing more simulations, we are going to pool everyone's data. On each of the three sticky notes you have, write one mean difference from the table above.

Add your sticky notes to the class dotplot on the board.

CHECKPOINT

10. How many sample mean difference fall above our observed mean difference?

11. What proportion of the randomized mean differences fall above our observed mean difference from #6?

12. What conclusion would you draw from this simulation analysis regarding the question of whether there is sufficient evidence that chunking aids memorization?

Our class results are actually a small subset of all the possible randomizations. To really investigate whether chunking aids memorization, we should perform 1000 trials and record the mean difference for each trial. Thankfully, we can use technology to quickly perform 1000 trials.

As you use technology to simulate the randomization, **<u>carefully read instructions</u>** -- they are going to instruct you how to use the applet as well as explain what is happening at each step.

Open the Randomization Test applet at http://www.rossmanchance.com/applets/randomization20/Randomization.html. Open the **Stacked Format** region of the applet. Note the order in which data is entered: the numeric value (number of letters memorized) is listed first followed by the grouping (CAT or CATF). Also, note that all members of group 1 (CAT) are listed first.

We need to enter the data collected in #3; list all CAT counts and labels first! Doing so will ensure that the CAT data is denoted by a red dot on the dotplot and the CATF data is denoted by a blue dot. Note that this is the same order that you used in the tactile simulation with red and blue slips of paper and a paper bag.

Once you are finished, press “OK”.

Note: If you ever want to return to the original data just click on the **Split** or **Stacked Format** buttons and then press OK.

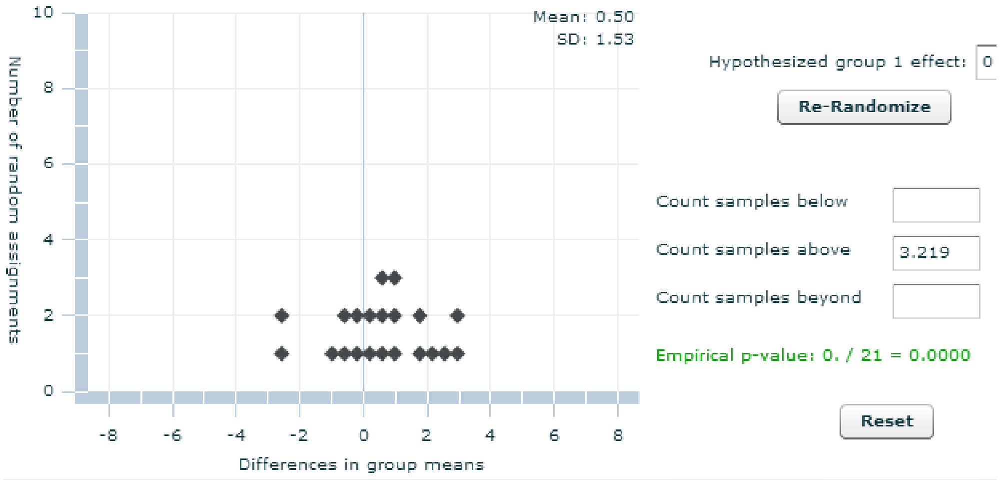
13. Our next step is to mix up all the data and randomly reassign it to the CAT or CATF group. Check **Animate**, and press the **Re-Randomize** button. Notice that the applet combines all the counts of letters memorized into one pile, mixes them up, and then redistributes them, at random, to the two groups. The applet also computes the mean of each group as well as the mean difference for you.

    What did you find for the difference in mean letters counted?

    $$\bar{x}_{\text{re-randomized } CAT} - \bar{x}_{\text{re-randomized } CATF} = \_\_\_\_\_\_\_\_\_\_\_\_$$

    Is this difference more extreme (farther from zero) than the observed mean difference in the data (from question #3)?

Now, we are going to quickly perform 20 trials. Change the **Number of repetitions** from 1 to 20. Select **Re-Randomize** again and observe how the dotplots of “could have been” data change. With each new “could have been” distribution (each new random assignment of counts to either CAT or CATF), the applet calculates the difference in group means and adds a dot to the dotplot at the bottom of the screen. You should see something like the display below.



Next, we are going to quickly perform 1000 trials. Press the **Reset** button in the bottom right corner of the applet. Change the **Number of repetitions** from 1 to 1000. Uncheck the **Animate** box. Select **Re-Randomize** and wait a few seconds.

Just as we did in our tactile simulation, the computer can count the number of mean differences that are at least as extreme as our observed difference (from #6).

14. Now enter the observed difference in means (question 6) into the **Count samples above** box and select the **Count** button.
    a. How many sample means fall above our observed mean difference?

    b. What proportion of the observations fall above our observed mean?

15. Interpret your value of p in terms of obtaining differences in the number of letters memorized as extreme as the data provide for each group.

16. What conclusion would you draw from this simulation analysis regarding the question of whether the number of letters memorized due to chunking is significantly less than that due to not chunking?

Student Handouts
Sample Solutions for the Teacher

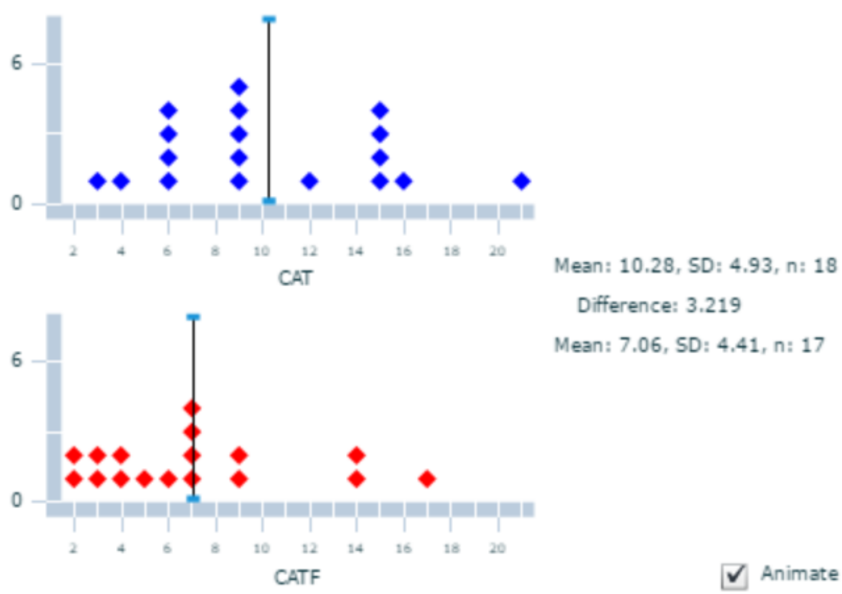
1. Some of the class had a sequence of letters that was organized into recognizable chunks of size 3, others had the same sequence of letters grouped into less recognizable chunks of 2, 3, or 4.

2. We would expect the group that had the CAT-FBI-USA grouping to be able to memorize more letters since the groups were recognizable. English language learners may find both groupings to be equally challenging.

3. Possible Data:

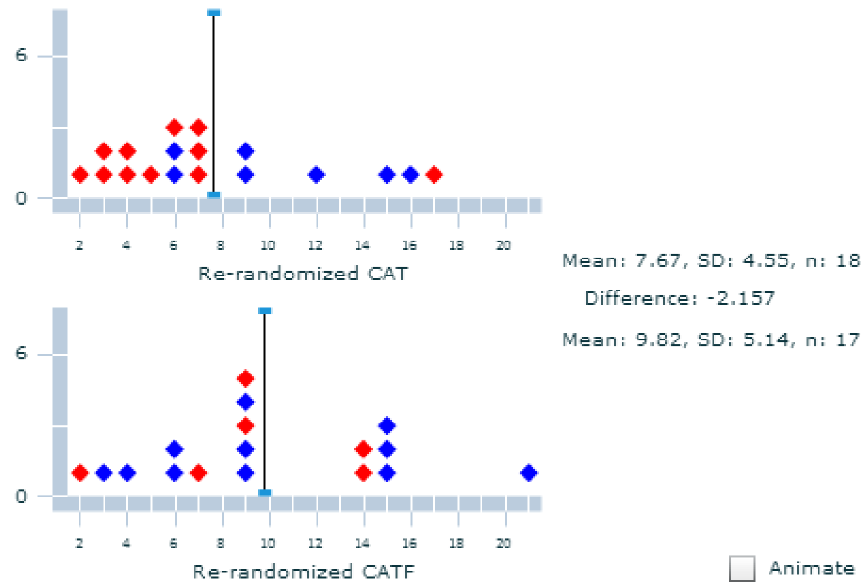
# Letters Correctly Recalled	CAT or CATF Group	# Letters Correctly Recalled	CAT or CATF Group	# Letters Correctly Recalled	CAT or CATF Group	# Letters Correctly Recalled	CAT or CATF Group
3	CAT	9	CAT	2	CATF	7	CATF
4	CAT	9	CAT	2	CATF	7	CATF
6	CAT	12	CAT	3	CATF	7	CATF
6	CAT	15	CAT	3	CATF	9	CATF
6	CAT	15	CAT	4	CATF	9	CATF
6	CAT	15	CAT	4	CATF	14	CATF
9	CAT	15	CAT	5	CATF	14	CATF
9	CAT	16	CAT	6	CATF	17	CATF
9	CAT	21	CAT	7	CATF		

4. Plots for the sample data.



5. Given the plots above the mean of the CAT sequence is larger than the CATF sequence. The data for CAT will likely be clumped at 3, 6, 9, 12, etc. letters and the CATF will likely occur at 3, 7, 10, 14, 17, etc. according to the grouping students were presented. For the CAT data, the center is about 10, the shape is trimodal (6, 9, 15), and most data falls between 6 and 15 letters. For the CATF data, the center is about 7, the data is nearly uniform with a single peak at 7, and most data falls between 2 and 9 letters.
6. $\bar{x}_{CAT} - \bar{x}_{CATF} = 3.219$
7. According to our data, it does appear that the CAT-FBI-USA aids memorization. The difference in the means is larger than 0.
   Simulation possible answer:
   g. $\bar{x}_A = \bar{x}_{\text{re-randomized } CAT} = 7.67 \qquad \bar{x}_B = \bar{x}_{\text{re-randomized } CATF} = 9.82$

$$\bar{x}_{\text{re-randomized } CAT} - \bar{x}_{\text{re-randomized } CATF} = -2.157$$

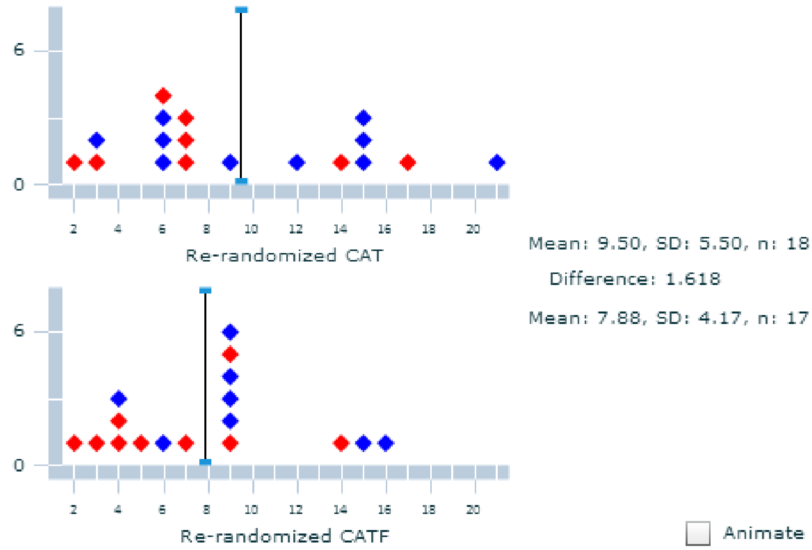


8. For the randomization above, the answer is no.
9. Possible mean differences:

	Trial 1	Trial 2	Trial 3
Mean difference: $\boldsymbol{\bar{x}_{\text{re-randomized } CAT} - \bar{x}_{\text{re-randomized } CATF}}$	-2.157	1.168	2.876

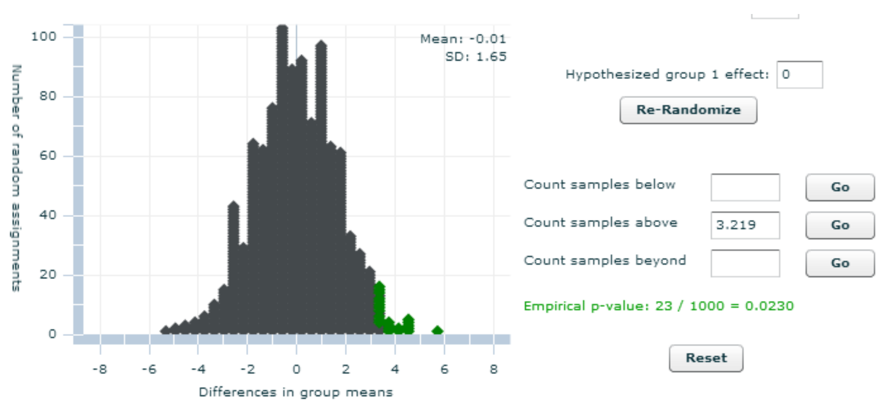
   Students should see pretty quickly that few trials actually returned a value as extreme (farther from zero) as the one we found in #6. However, this data should mirror the value of p that is found using technology, so it is possible that we obtain many results as extreme as the observed difference of means in #6.
10. Answers will vary, however the answer should be small if the final analysis shows the probability is small from the computed p value.
11. Answers should be consistent with part #10 and 3 times the number of groups in the class for the total number of randomizations.

12. If the probability is small, then we would conclude there is evidence that chunking aids memorization. If the probability is large, we have no evidence that chunking aids memorization.
13. Possible randomization for our sample data.



$\bar{x}_{\text{re-randomized } CAT} - \bar{x}_{\text{re-randomized } CATF} =$ ___1.618____
For the data given above, the answer is no, it's not more extreme.

Regardless of the data collected, the dotplot for 1000 repetitions should be roughly symmetric with mean approximately zero.



14.
   a. Answers will vary, however the answer should be small if the final analysis shows significance.
   b. In other words, what is the p-value? Answers should be consistent with part a. and 1000 total randomizations.
15. We expect that in many re-randomizations, this process will return approximately p (as a percentage) results that are as extreme as our answer to #6.
16. If p is small, then we would reject the statement that there is no evidence that chunking aids memorization in favor of the statement that there is a difference that chunking aids memorization. That is, we have significant evidence that chunking does aid memorization. If p is large, we have no evidence that chunking aids memorization.

Appendix
Sequence of letters arranged for CAT and CATF to use for Data Collection

CAT – FBI – USA – NFL – LOL – ABC – CBS – ACT – GPA – OMG – CPR

CATF – BIU – SAN – FLLO – LAB – CCB – SAC – TGP – AO – MGC – PR

CAT – FBI – USA – NFL – LOL – ABC – CBS – ACT – GPA – OMG – CPR

CATF – BIU – SAN – FLLO – LAB – CCB – SAC – TGP – AO – MGC – PR

CAT – FBI – USA – NFL – LOL – ABC – CBS – ACT – GPA – OMG – CPR

CATF – BIU – SAN – FLLO – LAB – CCB – SAC – TGP – AO – MGC – PR

CAT – FBI – USA – NFL – LOL – ABC – CBS – ACT – GPA – OMG – CPR

CATF – BIU – SAN – FLLO – LAB – CCB – SAC – TGP – AO – MGC – PR

CAT – FBI – USA – NFL – LOL – ABC – CBS – ACT – GPA – OMG – CPR

CATF – BIU – SAN – FLLO – LAB – CCB – SAC – TGP – AO – MGC – PR